TESIS – 142502

Evaluasi Metode *Hierarchical Clustering* Berbasis *Linkage* pada MWMOTE : Studi Kasus Data Akademik Universitas XYZ dan Data UCI

Meida Cahyo Untoro

NRP. 05111650010059

DOSEN PEMBIMBING

Prof. Dr. Ir. Joko Lianto Buliali, M.Sc.

NIP. 196707271992031002

PROGRAM MAGISTER

DEPARTEMEN INFORMATIKA

FAKULTAS TEKNOLOGI INFORMASI DAN KOMUNIKASI

INSTITUT TEKNOLOGI SEPULUH NOPEMBER

SURABAYA

2018

[Halaman ini sengaja dikosongkan]

Tesis disusun untuk memenuhi salah satu syarat memperoleh gelar
Magister Komputer (M.Kom.)
di
Institut Teknologi Sepuluh Nopember Surabaya

oleh:
Meida Cahyo Untoro
NRP. 05111650010059

Dengan judul :
Evaluasi Metode Hierarchical Clustering Berbasis Linkage pada MWMOTE :
Studi Kasus Data Akademik Universitas XYZ dan Data UCI

Tanggal Ujian : 26 Juli 2018
Periode Wisuda : September 2018

Disetujui oleh:

Prof. Dr. Ir. Joko Lianto Buliali, M.Sc.
NIP. 196707271992031002
....
(Pembimbing 1)

Dr. Eng. Nanik Suciati, S.Kom., M.Kom.
NIP. 197104281994122001
....
(Penguji 1)

Dr.Eng. Radityo Anggoro, S.Kom., M.Sc.
NIP. 198410162008121002
....
(Penguji 2)

Dr. Eng. Chastine Fatichah, S.Kom., M.Kom.
NIP. 197512202001122002
....
(Penguji 3)

KEMENTERIAN RISET, TEKNOLOGI, DAN PENDIDIKAN TINGGI
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
FAKULTAS TEKNOLOGI INFORMASI DAN KOMUNIKASI

Dekan Fakultas Teknologi Informasi dan Komunikasi,

Dr. Agus Zainal Arifin, S.Kom., M.Kom.
NIP. 197208091995121001

Evaluasi Metode *Hierarchical Clustering* Berbasis *Linkage* pada MWMOTE : Studi Kasus Data Akademik Universitas XYZ dan Data UCI

Nama : Meida Cahyo Untoro

NRP : 05111650010059

Pembimbing : Prof. Dr. Ir. Joko Lianto Buliali, M.Sc.

ABSTRAK

Ketidakseimbangan (*Imbalanced*) data terjadi pada berbagai macam data termasuk data akademik Universitas XYZ dan data UCI. Kasus tersebut menyebabkan adanya *misclassified* dikarenakan data mayoritas dominan terhadap data minoritas yang berakibat pada menurunnya nilai akurasi. Metode MWMOTE dapat menjadi pilihan dalam menyelesaikan kasus *imbalanced* melalui pembobotan dan *clustering*.

Penelitian ini bertujuan menangani permasalahan *imbalanced* dataset akademik di Universitas XYZ angkatan 2014 dan 2015 dan data UCI dengan mengevaluasi *hierarchical clustering.* Tujuan tersebut dicapai dengan mengevaluasi tiga metoda *hierarchical cluster* sebagai salah satu sub proses pada MWMOTE untuk menghasilkan data sintetik yang lebih representatif.

Hasil yang didapat dari penelitian ini adalah ketiga metoda AHC tersebut tidak memberikan perbedaan yang signifikan dalam perbaikan akurasi MWMOTE pada data akademik dan 7 data UCI yang diuji dengan *one*-way ANOVA dengan nilai sig/*alpha* > 0.05.

Keyword : *Imbalanced*, *Hierarchical Clustering*, *Linkage*, MWMOTE, Data Sintetik

[Halaman ini sengaja dikosongkan]

Evaluation of Linkage-Based Hierarchical Clustering Method on MWMOTE : Case Study on Data Academic of XYZ and Data UCI

Name : Meida Cahyo Untoro
Student identity Number : 05111650010059
Supervisor : Prof. Dr. Ir. Joko Lianto Buliali, M.Sc.

ABSTRACT

Imbalanced data occurs in a variety of data including academic data of XYZ University and UCI data. The case is misclassified due to the dominant majority of data on minority data resulting in a decrease in accuracy. The MWMOTE method can be an option in solving imbalanced cases through weighting and clustering. This study aims to address the academic incremental dataset issues at XYZ University force 2014 and 2015 and UCI data by evaluating hierarchical clustering.

This objective was achieved by evaluating three hierarchical cluster methods as one of the sub processes on MWMOTE to produce more representative synthetic data.

The results obtained from this study are the three AHC methods do not give significant difference in MWMOTE accuracy improvement on academic data and 7 UCI data tested with one-way ANOVA with sig / alpha value > 0.05.

Keyword : Imbalanced, Hierarchical Clustering, Linkage, MWMOTE, Data Synthetic.

[Halaman ini sengaja dikosongkan]

KATA PENGANTAR

Puji syukur atas kehadirat Allah SWT, karena atas limpahan dan karunia-Nya lah penulis dapat menyelesaikan buku tesis penelitian ini dengan baik dan lancar. Terima kasih juga tidak lupa penulis ucapkan kepada seluruh pihak yang telah membantu penulis dalam penyusunan buku tesis ini, tanpa bantuan dari seluruh pihak yang telah berkontribusi dalam penyusunan buku tesis ini mungkin buku ini tidak akan selesai.

Pada kesempatan ini, penulis menyampaikan ucapan terima kasih yang sebesar-besarnya kepada:

1. Allah SWT atas limpahan rahmat, karunia, serta ilmu-Nya sehingga penulis dapat menyelesaikan tesis ini dengan baik.
2. Kedua orang tua penulis yang selalu memberi dukungan moril, materil, dan selalu mendoakan penulis untuk bisa menyelesaikan studi.
3. Bapak Prof. Dr. Ir. Joko Lianto Buliali, M.Sc. selaku dosen pembimbing penulis yang telah mempercayai saya menjadi anak bimbing beliau, memberikan bantuan dan motivasi pada penulis sehingga dapat menyelesaikan buku tesis ini.
4. Ibu Dr. Eng. Nanik Suciati, S.Kom., M.Kom., Ibu Dr.Eng. Chastine Fatichah S.Kom. M.Kom., dan Bapak Dr.Eng. Radityo Anggoro, S.Kom., M.Sc. selaku dosen penguji yang telah memberi saran, masukan, dan koreksi dalam tesis ini.
5. Ibu Mastuti Widianingsih, S.Si.,M.Sc. selaku pembimbing tidak tertulis dan partner hidup, yang memberikan segala kemampuan dan waktu sehingga penulis dapat menyelesaikan tesis ini.
6. Bapak Endyk Noviyantono, Bapak Mudjahidin, Bapak Mohammad Yazdi Pusadan, Bapak Tora Fahrudin, Ibu Eviana Tjatur Putri, Ibu Anggreni, Ibu Myrna Ermawati, Ibu Eva Firdayanti Bisono, Bapak Achmad Saiful, dan Bapak Reza Prasetya Prayogo sebagai keluarga ABJ yang banyak sekali memberikan kritik dan saran selama pengerjaan tesis.
7. Rekan-rekan seangkatan 2016 S2 Informatika ITS yang telah menemani penulis selama 2 tahun dalam penempuh studi magister.
8. Serta pihak-pihak lain yang tidak dapat disebutkan satu persatu.

Penulis menyadari bahwa tesis ini masih jauh dari kesempurnaan dan banyak kekurangan. Untuk itu dengan segala kerendahan hati penulis mengharapkan kritik dan saran yang membangun dari pembaca.

Surabaya, 26 Juli 2018

Meida Cahyo Untoro

DAFTAR ISI

DAFTAR GAMBAR

[Halaman ini sengaja dikosongkan]

DAFTAR TABEL

[Halaman ini sengaja dikosongkan]

BAB 1
PENDAHULUAN

1.1 Latar Belakang

Universitas XYZ memiliki data akademik yang terus bertambah setiap tahun. Data akademik terdiri dari data presensi, data nilai (Tugas, Quis, UTS, ataupun UAS), dan data aktivitas lainnya yang diakumulasi setiap bulan dalam 1 semester. Akumulasi data tersebut dapat menjadi tolak ukur keberhasilan proses akademik seorang mahasiswa maupun institusi. Seringkali terjadi ketidakseimbangan data, misalnya jumlah mahasiswa lulus tahun pertama secara normal lebih banyak dari pada tidak lulus. Dalam sistem komputasi, ketidakseimbangan data disebut dengan *imbalanced* (Chawla *et al.*, 2002; Barua *et al.*, 2014; Fahrudin *et al.*, 2016). Salah satu permasalahan yang terjadi dalam kasus mahasiswa tidak lulus, karena setiap mahasiswa memiliki tahapan dalam pembelajaran.

Pada penelitian ini data akademik didapat dari penelitian "Deteksi Dini Kegagalan Akademik Serta Multilabelisasi Permasalahan Mahasiswa Dari Data Media Sosial" (Fahrudin, 2017). Data akademik yang didapat bersifat *imbalanced*, dimana mahasiswa lulus tahun pertama secara normal lebih banyak dari pada tidak lulus. Adanya permasalahan data *imbalanced* antara mahasiswa yang tidak dapat melewati evaluasi tahap pertama dalam waktu normal (1 tahun) (data minoritas). Mahasiswa mampu melewati tahap pertama (data mayoritas), menyebabkan akurasi data minoritas menjadi rendah (Jayasree and Gavya, 2015). Pendistribusian yang tidak seimbang (*class imbalanced*) menimbulkan kejadian klasifikasi yang lebih condong ke jumlah data mayoritas (*negative*) dibandingkan dengan jumlah data minoritas (*positive*) (Sáez *et al*., 2016).

Mellor *et al.* (2015) menyatakan kasus *misclassified* diakibatkan oleh *imbalanced* dataset. Kasus *imbalanced* dapat mengelompokkan data menjadi 2, yaitu data minoritas dan mayoritas (Seiffert, Khoshgoftaar and Van Hulse, 2009; Liu W, Chawla, 2010; Barua *et al.*, 2014). Selain itu, *imbalanced* dapat menyebabkan pembuatan model yang buruk (Gong and Kim, 2017) serta *overfitting* dan penurunan akurasi klasifikasi (Fakhruzi, 2018). Berdasarkan permasalahan tersebut diperlukan penanganan *imbalanced* dataset, sehingga didapat model yang memiliki ketepatan prediksi yang maksimal pada semua kelas data (Rivera, 2017; Piri, Delen and Liu, 2018). Kasus *imbalanced* dapat ditangani dengan metode *oversampling* (Chawla *et al.*, 2002; Barua *et al.*, 2014; Piri *et al.*, 2018) dan *undersampling* (Ng *et al.*, 2015; Purwar and Singh, 2015; John and Jayasudha, 2017). *Oversampling* dilakukan dengan membuat replika (*resample*) data minoritas, sedangkan *undersampling*, mengurangi data mayoritas sehingga didapatkan data mayoritas dan minoritas seimbang (Barua *et al.*, 2014; Ng *et al.*, 2015; Fahrudin *et al.*, 2016; John and Jayasudha, 2017; Fakhruzi, 2018).

Metode *oversampling* berlebihan dapat menyebabkan *overfitting* dan *undersampling* berlebihan dapat berpengaruh pada hilangnya beberapa informasi penting yang terdapat pada dataset (Seiffert *et al.*, 2009; Napierała, 2012; Ma and Fan, 2017). *Synthethic Minority Oversampling Technique* (SMOTE) dapat menangani *overfitting* data sintetik melalui pendekatan k NN (Chawla *et al.*, 2002; El-Sayed *et al.*, 2016) dengan penggunaan variabel (Blagus *et al.*, 2013). Barua *et al.*, 2014 mengusulkan *Majority Weighted Minority Oversampling Technique* (MWMOTE) sebagai metode perbaikan dari SMOTE melalui pembuatan data sintetik dengan pembobotan dan *clustering* data minoritas. Hasil usulan tersebut ternyata mampu menurun derajat bias atau *noise* serta menghasilkan data sintetik dengan tingkat akurasi lebih baik dari proses *clustering*.

Peningkatan akurasi *classifier* pada data *imabalance* dapat ditangani dengan menggunakan *agglomerative hierarchy clustering* (AHC). AHC memiliki kinerja yang akurat bahkan ketika *cluster* memiliki rasio *imbalanced* rendah pada data minoritas serta terjadi tumpang tindih data mayoritas dengan cara partisi dan deteksi *outlier* (Beyan and Fisher, 2015). Li, Wang and Hao, 2009 melakukan penelitian terkait c*omparison of Cluster Ensembles* pada *average*, *complete*, dan *single linkage* didapat peningkatan performa akurasi 10 dataset UCI. AHC pada pengenalan wajah berbasis (*Content Based Image Retrieval*) CBIR dapat meningkatkan performa komputasi dan dilakukan uji validitas *cluster* (*cophenetic correlation coefficient / CCC*) (Fachrurrozi *et al.*, 2017). Hasil analisis *cluster* AHC berdasarkan dendogram memberikan hasil yang berbeda-beda setiap dataset uji coba (Hakim, Subanar and Winarko, 2010).

Data sintetik penting dalam pembuatan pola data minoritas. Permasalahan muncul ketika data yang tidak seimbang menyebabkan kelompok tertentu menjadi *misclassified*. Penelitian ini bertujuan untuk menangani permasalahan data tidak seimbang (*imbalanced*) dengan membuat data sintetik menggunakan MWMOTE dan mengevaluasi metode *cluster* berbasis *linkage*. Pengukuran performa pada model yang dihasilkan berdasarkan pada nilai *cophenetic correlation coefficient*, F-*measure*, *recall*, *precision*, dan *accuracy*.

1.2 Perumusan Masalah

Rumusan masalah yang diangkat dalam penelitian ini adalah sebagai berikut.

1. Bagaimana cara menanganani *imbalanced* dataset akademik dan dataset UCI kelas minoritas dengan metode MWMOTE.
2. Bagaimana cara mengevaluasi metode *hierarchical clustering* berbasis *linkage* pada MWMOTE.

1.3 Tujuan Penelitian

Tujuan yang akan dicapai dalam penelitian ini adalah

1. Menangani permasalahan *imbalanced* data akademik Universitas XYZ angkatan 2014 dan 2015 dan dataset UCI.
2. Menganalisis penggunaan metode *hierarchical clustering* pada algoritma MWMOTE dalam perbaikan kualitas *cluster* dan akurasi.

1.4 Manfaat Penelitian

Manfaat dari penelitian ini adalah memberikan solusi penanganan data *imbalanced* pada data akademik Universitas XYZ dan dataset UCI dengan pengujian *hierarchical clustering* metode *cluster* pada algoritma MWMOTE.

1.5 Kontribusi Penelitian

Kontribusi dalam penelitian ini adalah mengevaluasi metode *Hierarchical Clustering* pada proses *cluster* MWMOTE pada kualitas *cluster* dan perbaikan akurasi klasifikasi pada data akademik Universitas XYZ angkatan 2014 dan 2015 dan data UCI.

1.6 Batasan Penelitian

Penelitian ini dibatasi pada hal-hal sebagai berikut:

1. Data yang digunakan merupakan data akademik Universitas XYZ selama 1 semester angkatan 2014 dan 2015 (Bulan Agustus – Desember).
2. Data UCI yang digunakan sebanyak 10 dataset, antara lain : Abalone, Breast, Ecoli, Glass, Libra, OCR, Robot, Satimage, Wine dan Yeast.
3. *Cluster* yang digunakan *average*, *complete*, dan *single linkage* dan dievaluasi dengan *cophenetic correlation coefficient*.
4. *Classifer disscion tree* (J48) digunakan untuk mengevaluasi metode *oversampling* dengan nilai parameter *precision*, *recall*, F-*measure* dan *accuracy*.

BAB 2
KAJIAN PUSTAKA

2.1 Data Mining

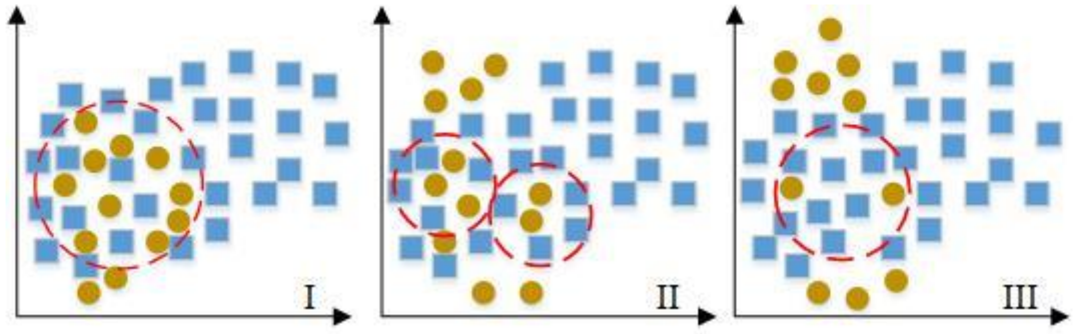
Data mining sering juga disebut sebagai *knowledge discovery in* data*base* (KDD) yaitu proses ekstraksi dan analisis sejumlah data. Proses tersebut bertujuan untuk *summarization* (generalisasi data), klasifikasi, *clustering, association,* dan prediksi sebagai basis pengetahuan untuk keperluan pengambilan keputusan. Data mining memiliki beberapa tahapan, data *cleaning*, data *integration*, data *selection*, data *transformation*, *pattern evaluation*, dan *knowledge presentation* (Ristoski and Paulheim, 2016).

Pengenalan pola diperlukan dalam mempelajari klasifikasi data berdasarkan kelas atau kategori dan mengetahui label data. Data yang digunakan bisa berupa gambar, *signal*, subjek manusia (misalnya pasien, mahasiswa), ataupun pengukuran lain yang perlu klasifikasi (Fahrudin *et al.*, 2017). *Machine learning* menjadi salah satu metode analisis dalam data mining. Hasil analisis metode tersebut dapat berupa deskripsi, prediksi, klasifikasi, *clustering*, dan asosiasi (Larose and Larose, 2014).

2.2 *Imbalanced*

Imbalanced adalah ketidakseimbangan data sehingga didapatkan kelas mayoritas dan kelas minoritas (Guo *et al.*, 2016). Ketidakseimbangan dataset menjadi permasalahan yang dihadapi oleh para peneliti dalam domain *imbalanced*. Kompleksitas data yang diperoleh para peneliti memiliki 3 sifat, yaitu *overlap*, *small disjunct* dan *outlier* (**Gambar 2.1**) (Mahmood, 2015).

Klasifikasi *imbalanced* menyebabkan *misclassification* sehingga nilai akurasi yang dihasilkan cenderung buruk serta memungkinkan kelas minoritas dianggap sebagai *outlier* (Seiffert *et al.*, 2009). *Imbalanced* umumnya dapat ditangani dengan 3 cara, yaitu *Random Under-Sampling* (RUS), *Random Over-Sampling* (ROS), dan *Hybrid Sampling Methods* (HSM) (Mahmood, 2015).



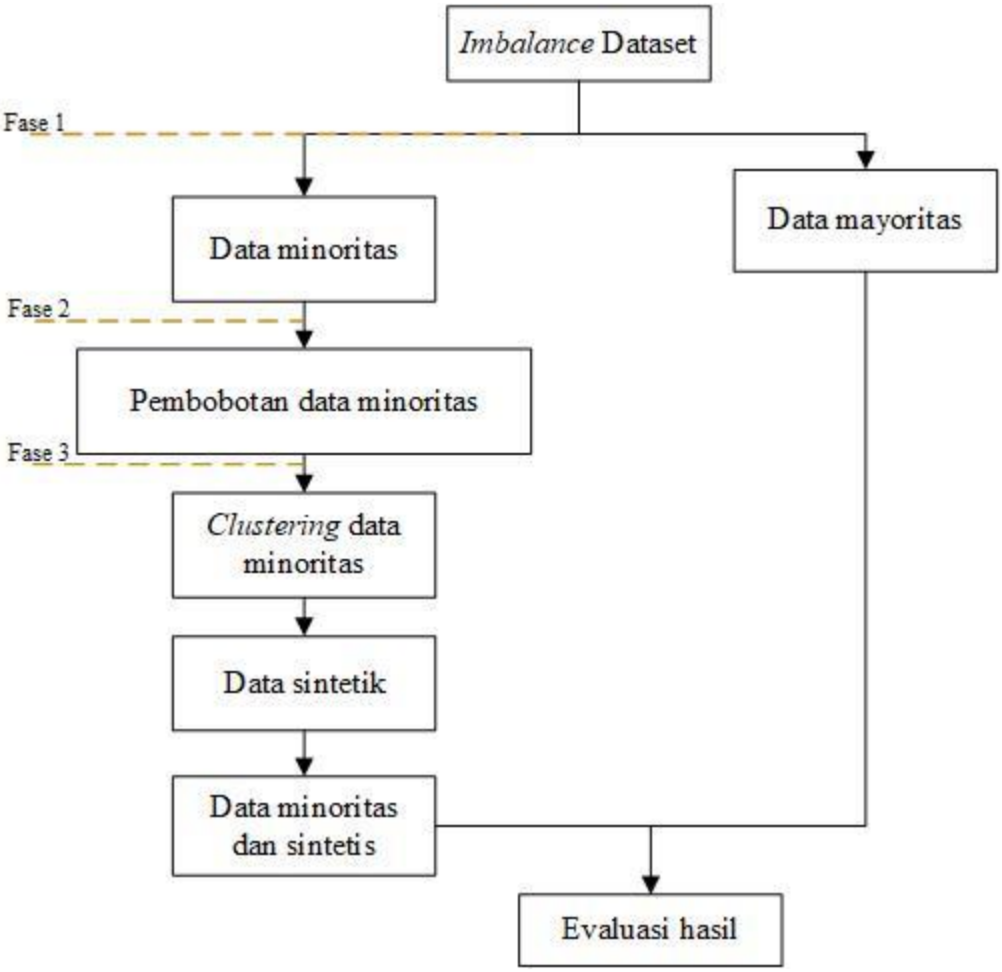
Gambar 2.1 *Overlap* (I), *Small disjunct* (II), *Outlier* (III)

Random Under-Sampling (RUS) berguna menghitung selisih antara kelas mayoritas dan minoritas (sebagai langkah pertama) kemudian dilakukan perulangan selisih hasil perhitungan. Cara RUS lebih efektif dan cepat dalam proses prediksi *imbalanced class*, misalnya pada kasus *software defects* (Irawan *et al.*, 2015). *Random Over-Sampling* (ROS), menyeimbangkan distribusi kelas minoritas dan kelas mayoritas. Dalam kasus ini, dilakukan replikasi acak kelas minoritas sampai didaptkan jumlah kelas minoritas sama dengan kelas mayoritas. Penelitian menunjukkan penggunaan ROS dapat mengurangi terjadinya *overfitting* (Zheng, Cai and Li, 2015). *Hybrid Sampling Methods* (HSM) merupakan penggambungan metode RUS dan ROS (Seiffert *et al.*, 2009). Kelas minoritas pada umumnya menjadi acuan prediksi yang bersifat negatif (*noise*), sehingga mampu menurunkan kinerja model kelas mayoritas (Napierała, 2012).

2.3 *Oversampling*

Oversampling menjadi salah satu cara dalam menangani *imbalanced* dataset dengan cara pembuatan data sintetik kelas minoritas sesuai jumlah selisih kelas mayoritas dan minoritas. Pembuatan data sintetik dapat dilakukan dengan metode *Synthetic Minority Oversampling Technique* (SMOTE). Pembuatan data sintetik kelas minoritas bertujuan menyeimbangkan rasio antar kelas minoritas dengan mayoritas. Data tersebut dibuat berdasarkan kelas minoritas dengan prinsip K-*Nearest Neighbor* (Chawla *et al.*, 2002).

Prinsip K-*Nearest Neighbor* digunakan untuk mencari nilai terdekat kelas minoritas pada *Majority Weighted Minority Oversampling Technique* (MWMOTE). Pembuatan data sintetik pada MWMOTE terdapat 3 tahap, yaitu identifikasi sampel kelas minoritas dan kelas mayoritas pada dataset, pembobotan kelas minoritas, dan *clustering* (Barua *et al.*, 2014) (**Gambar 2.2**).



Gambar 2.2 Algoritma MWMOTE (Barua *et al.*, 2014)

Pembuatan data sintetik MWMOTE memiliki cara kerja dengan memilih sampel dengan memisahkan data kelas mayoritas dan kelas minoritas. Tahap selanjutnya dilakukan proses pembobotan untuk mengetahui data kelas minoritas yang mendekati *borderline.* Pembobotan diberikan berdasarkan kedekatan data dengan *borderline*, jumlah anggota *cluster* kecil, dan *cluster* kelas minoritas (berjumlah banyak) yang berada dalam kelas mayoritas. *Clustering* menjadi proses lanjutan tahap pembobotan yang dilakukan dengan *algomerative hiearacal clustering* sehingga didapatkan data sintetik (Barua *et al.*, 2014).

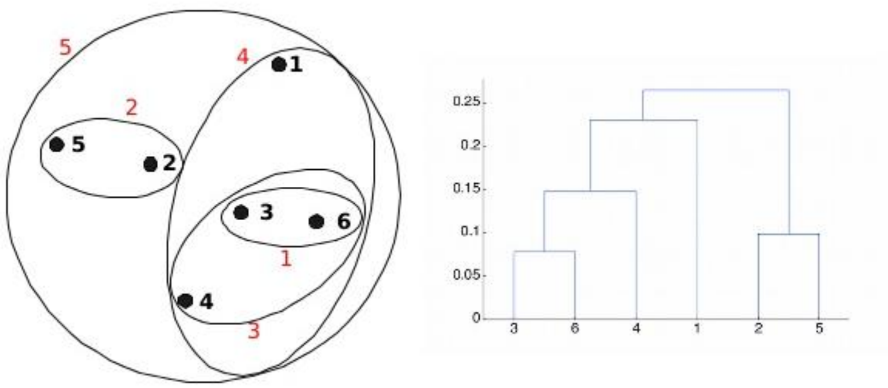
2.4 Analisis *Cluster*

Analisis *cluster* merupakan teknik analisis statistik yang ditujukan untuk menempatkan sekumpulan data ke dalam dua atau lebih *cluster* berdasarkan kemiripan data berdasarkan variasi atribut (Köhn and Hubert, 2015). Secara umum metode *clustering* ada dua, yaitu *hierarchical clustering* dan *non-hierarchical clustering*.

Hierarchical clustering dilakukan melalui proses pengelompokkan *clutering* yang berasal dari *cluster* individu yang memiliki kemiripan. Ada 2 macam *hierarchical clustering* yaitu *agglomerative* dan *devisive clustering.* Hasil akhir *agglomerative clustering* akan didapatkan satu *cluster* utuh yang berisi semua data, sedangkan pada *devisive clustering* menggunakan prinsip *top bottom* (1 *cluster* dipisahkan dalam 2 ataupun 3 *cluster* yang berbeda) (Köhn and Hubert, 2015). *Hierarchical agglomerative* terdiri dari *linkage, Ward's,* dan *centroid method* (Lin *et al.*, 2017). *Linkage method* terdiri dari 3 bagian, yaitu :

2.4.1 *Average linkage*

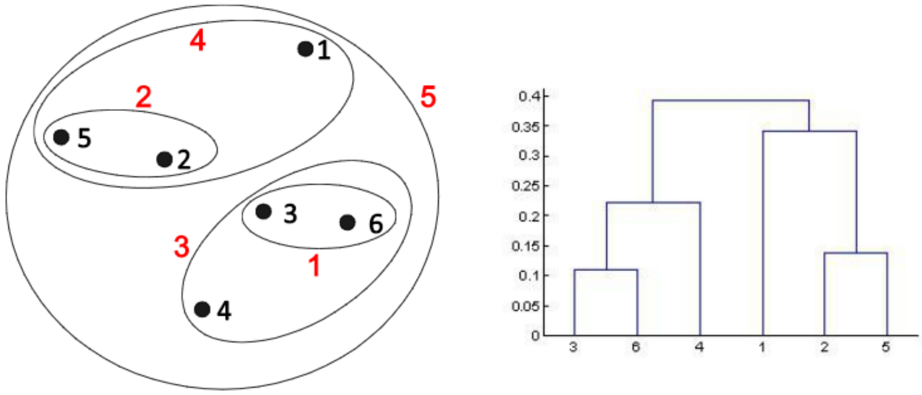
Jarak antara dua *cluster* didefinisikan sebagai jarak rata-rata antara setiap titik dalam satu *cluster* ke setiap titik pada *cluster* lain. Prosedur *average linkage* dengan mendefinisikan matrik D untuk memperoleh data paling dekat, sebagai contoh **Gambar 2.3**. Data U dan V digabungkan ke dalam bentuk *cluster* (UV) (Huth, Nemesova and Klimperova, 1993), sehingga didapatkan jarak rata-rata antara UV ataupun jarak antar *cluster* yang lainnya.



Gambar 2.3 *Average linkage*

2.4.2 *Complete linkage*

Complete linkage merupakan jarak maksimum antar dua *cluster* (**Gambar 2.4**). Prinsip jarak yang digunakan adalah jarak terjauh antar data (Pusadan, M Y, Buliali, J L, Ginardi, 2016).

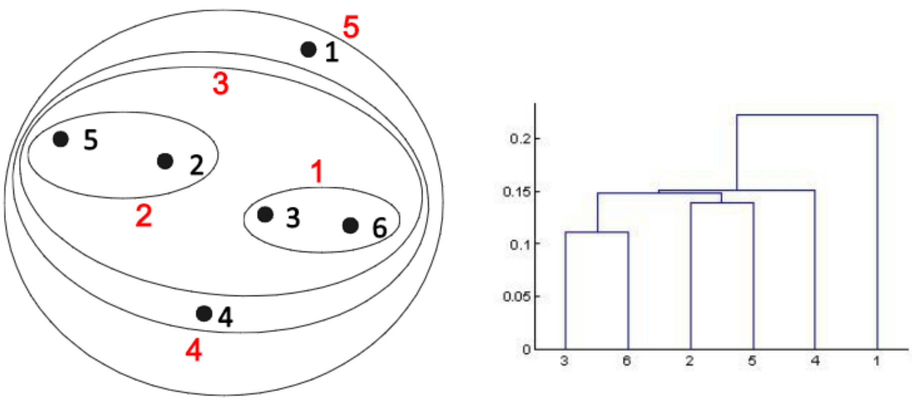


Gambar 2.4 *Complete linkage*.

$Dist(u[i]$ dan $dist(v[j]$ masing-masing adalah jarak antara anggota yang paling jauh dari cluster U dan V serta cluster I dan K (Abd Rahman, Abu Bakar and Zulkefli, 2016).

2.4.3 *Single linkage*

Metode ini menggunakan prinsip jarak minimum antar dua data terdekat yang akan membentuk *cluster* pertama (**Gambar 2.5**). Langkah selanjutnya memiliki 2 kemungkinan, data ketiga akan bergabung dengan *cluster* yang telah terbentuk atau data ketiga akan bergabung dengan data yang lain untuk membentuk *cluster* baru (Almeida *et al.*, 2007).



Gambar 2.5 *Single linkage*

2.5 *Cophenetic Correlation Coefficient*

Pengujian hasil *clustering* dilakukan dengan validitas *cluster*. Uji validitas *cluster* diperlukan untuk menentukan kualitas hasil analisis *cluster*. Salah satu ukuran yang dapat digunakan untuk menguji validitas hasil *cluster* metode hirarki adalah dengan *cophenetic correlation coefficient* (CCC). *Cophenetic Correlation Coefficient* merupakan nilai korelasi antara jarak *euclidean* dan *dendrogram* (*cophenetic matrix*) (Kumar and Toshniwal, 2016). Nilai CCC dikatakan baik apabila nilai *cluster* mendekati 1 dan buruk jika mendekati nilai negatif 1 (Bouguettaya *et al.*, 2015, Kumar and Toshniwal, 2016).

2.6 Klasifikasi *Decision Tree*

Klasifikasi adalah proses untuk menemukan model atau fungsi yang menjelaskan atau membedakan konsep atau kelas data dengan tujuan memperkirakan kelas yang tidak diketahui dari suatu data. Klasifikasi data melibatkan dua proses, yaitu proses *training* dan *testing.* Proses *training* digunakan dalam pembuatan model atau pola, sedangkan proses *testing* digunakan untuk mengetahui keakuratan model hasil *training* (Fahrudin, Buliali and Fatichah, 2017). *Decision tree* adalah algoritma yang paling banyak digunakan untuk masalah klasifikasi. *Decision tree* terdiri dari beberapa simpul yaitu *tree's root*, *internal node* dan *leafs*. Konsep entropi digunakan untuk penentuan pada atribut mana sebuah pohon akan terbagi (*split*) (Lesmana, 2012). *Decision tree* dikatakan baik jika mendekati nilai model yang terbentuk dan diukur nilai *presicion, recall, F-measure,* dan *accuracy* (Fahrudin, Buliali and Fatichah, 2017).

2.7 Uji Statistik

Uji statistik digunakan untuk menguji perbedaan antar grup, kelompok dan jenis perlakuan pada sebuah penelitian. Uji statistik ada dua macam parametrik dapat menggunakan *Analysis of Variance* (ANOVA) dan Friedman merupakan non parametrik. ANOVA digunakan untuk menganalisis perbedaan rata-rata antara kelompok dan prosedur terkait (seperti "variasi" antara kelompok) (Gregory, 2009) (**Tabel 2.1**). **Tabel 2.2** adalah gambaran hasil dari uji statistik menggunakan ANOVA. Dimana banyaknya grup, kelompok, dan perlakuan dinyatakan dalam “k”, “n” adalah jumlah data dari grup, kelompok dan perlakuan $\sum_{i-1}^{k} n_i$.

Tabel 2.1 Tabel ANOVA

Sumber Variasi	Derajat bebas	*Sum of square*	*Mean square*	F_{hit}
Perlakuan	(k – 1)	*Sum of square Perlakuan*	$MSP = A = \frac{SSP}{(k-1)}$	$\frac{A}{B}$
Error	(n – k)	*Sum of square Error*	$MSP = B = \frac{SSE}{(n-k)}$	
Total	(n – 1)	*Sum of square* Total		

Tabel 2.2 Hasil Evaluasi *One Way* ANOVA

	Sum of Squares	Df	*Mean Square*	F	Sig.
Between Groups	490.433	2	245.217	91.866	.000
Within Groups	152.150	57	2.669		
Total	642.583	59			

ANOVA dibagi menjadi beberapa bagian salah satunya *One Way* ANOVA, *two-way* ANOVA. *One Way* ANOVA bertujuan untuk membandingkan lebih dari dua rata-rata, sedangkan gunanya untuk menguji kemampuan generalisasi. Maksudnya dari signifikansi hasil penelitian, jika terbukti berbeda berarti kedua sampel tersebut dapat digeneralisasikan (data sampel dianggap dapat mewakili populasi). Prosedur dalam *One Way* ANOVA harus berdistribusi normal dan bervarian *homogeny*. Serta pembuatan hipotesis yang menyatakan tidak adanya hubungan antara variabel independen (X) dan variabel dependen (Y) disebut H0. H1 adalah hipotesis yang menyatakan adanya hubungan antara variabel independen (X) dan variabel dependen (Y) yang diteliti. Tujuan dari hipotesis berfungsi memusatkan peneliti pada focus permasalahan yang diselesaikan.

Jika H0 diterima berarti semua perlakuan yang dicobakan memberikan pengaruh yang sama, tetapi jika H1 yang diterima berarti paling sedikit terdapat sepasang nilai tengah perlakuan yang berbeda. Untuk mengetahui pasangan perlakuan mana yang mempunyai nilai tengah yang berbeda tersebut, maka perlu dilakukan pengujian lanjutan untuk mengetahui perbedaan diantara nilai tengah perlakuan tersebut. Pengujian tersebut diistilahkan dengan uji lanjutan atau biasa juga disebut uji pembanding berganda.

Penggunaan uji lanjutan digunakan untuk mengetahui pasangan perlakuan mana yang mempunyai nilai tengah yang berbeda. Untuk menentukan jenis uji lanjutan yang sesuai maka harus diperhatikan apakah uji yang akan digunakan adalah untuk perbandingan yang bersifat terencana atau tidak. Perbandingan terencana adalah perbandingan yang memang direncanakan sebelum data suatu percobaan diperoleh atau sebelum percobaan dilakukan, sedangkan perbandingan tidak terencana adalah perbandingan yang dilakukan setelah data diperoleh.

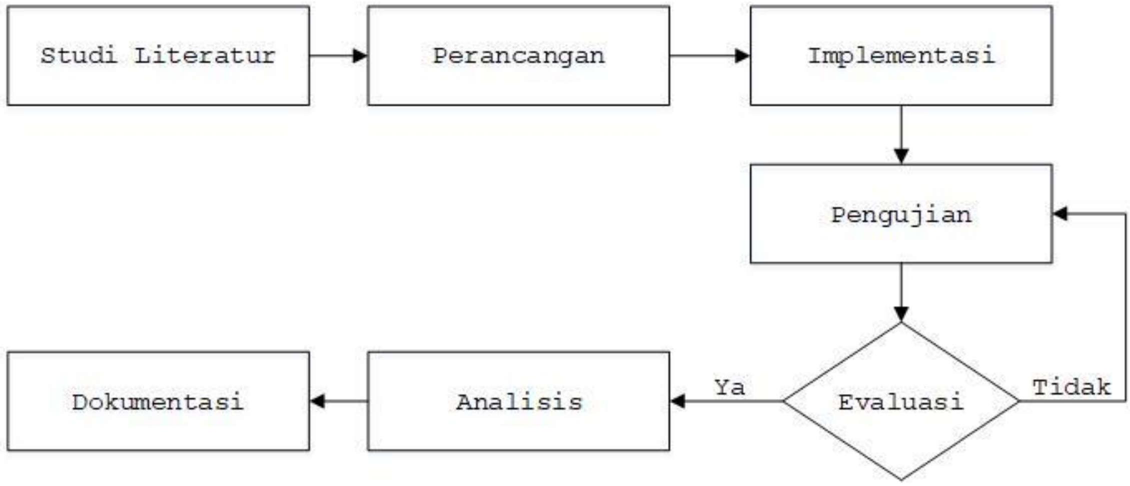
Uji Duncan adalah prosedur perbandingan dari nilai tengah perlakuan (rata-rata perlakuan) untuk semua pasangan perlakuan yang ada. Uji lanjut ini menggunakan nilai pembanding sebagai alat uji sesuai dengan jumlah nilai tengah atau rataan yang ada di wilayah dua perlakuan yang dibandingkan. Uji Duncan juga digunakan untuk melihat adanya pengaruh antar perlakuan yang diuji Duncan atau juga dikenal dengan istilah *Duncan Multiple Range Test* (DMRT) memiliki nilai kritis yang tidak tunggak tetapi mengikuti urutan rata-rata yang dibandingkan (**Tabel 2.3**). Nilai kritis uji Duncan dinyatakan dalam nilai *least significant range*.

Tabel 2.3 Hasil Uji Lanjut Duncan

Metode	N	*Subset for alpha* = 0.05		
		1	2	3
Metode B	20	66.50000		
Metode A	20		68.45000	
Metode C	20			**73.30000**
Sig.		1.000	1.000	1.000

BAB 3
METODOLOGI PENELITIAN

Bab 3 berisi pemaparan proses penelitian yang diawali dengan studi literatur (**Gambar 3.1**).



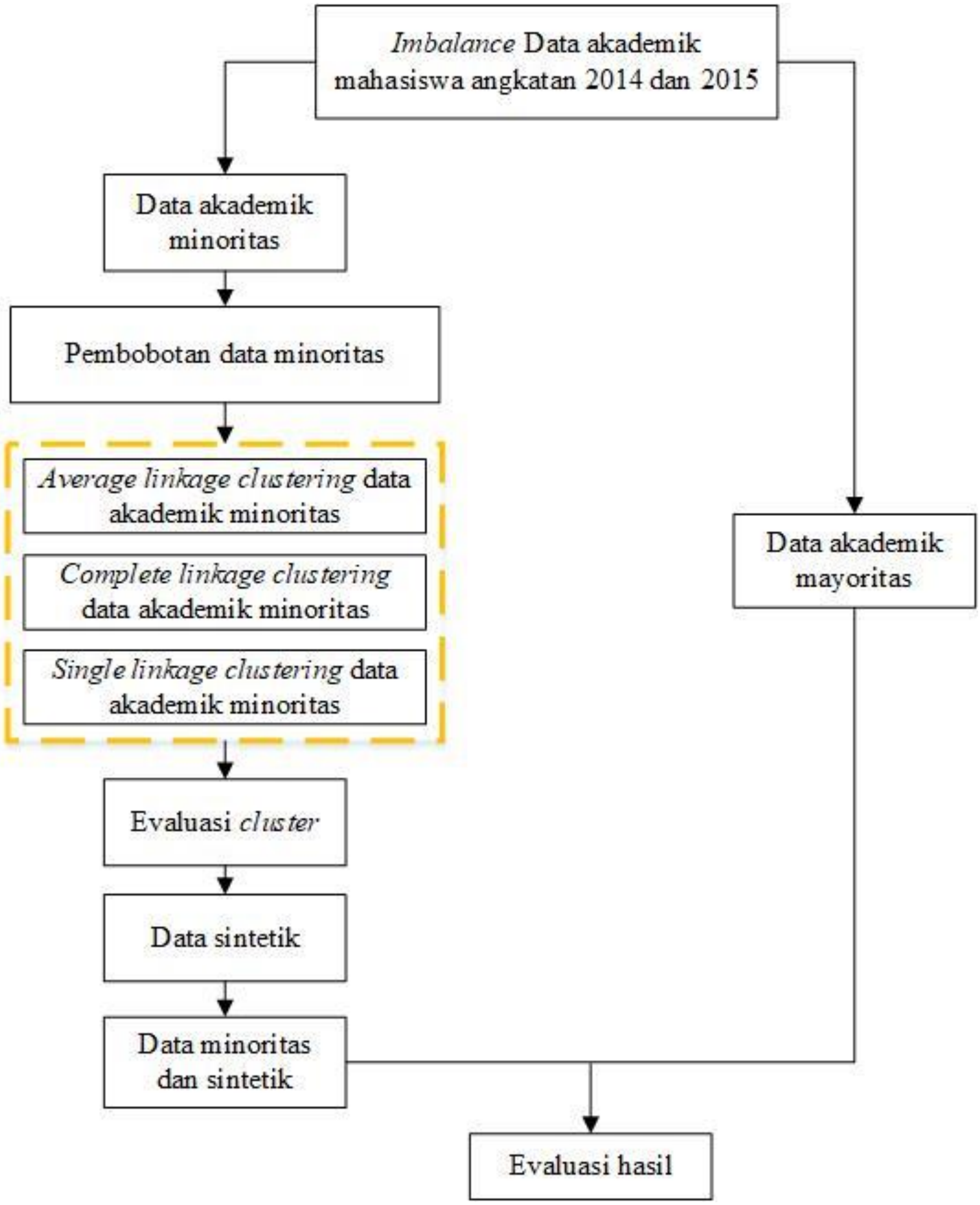
Gambar 3.1 Diagram Blok Metode Penelitian

3.1 Studi Literatur

Penelitian diawali dengan proses pengkajian yang berkaitan dengan topik penelitian. Pada penelitian ini, referensi yang digunakan diperoleh dari jurnal yang memiliki hubungan dengan data akademik, data *mining*, *imbalanced* dataset, *Agglomerative Hierarchical Clustering* (AHC), *oversampling*, *undersampling*, dan *Majority Weighted Minority Over-sampling Technique* (MWMOTE).

3.2 Perancangan

Penelitian dilakukan dengan menggunakan data akademik mahasiswa angkatan 2014 dan 2015 dalam 1 semester. Metode yang digunakan adalah *Majority Weighted Minority Over-sampling Technique* (MWMOTE) dengan mekanisme *synthetic oversampling techniqu* (Chawla *et al.*, 2002, Barua *et al.*, 2014, Jayasree and Gavya, 2015, Fahrudin, Buliali and Fatichah, 2016, Ma and Fan, 2017, Piri, Delen and Liu, 2018). Perancangan penelitian tersaji pada **Gambar 3.2**.

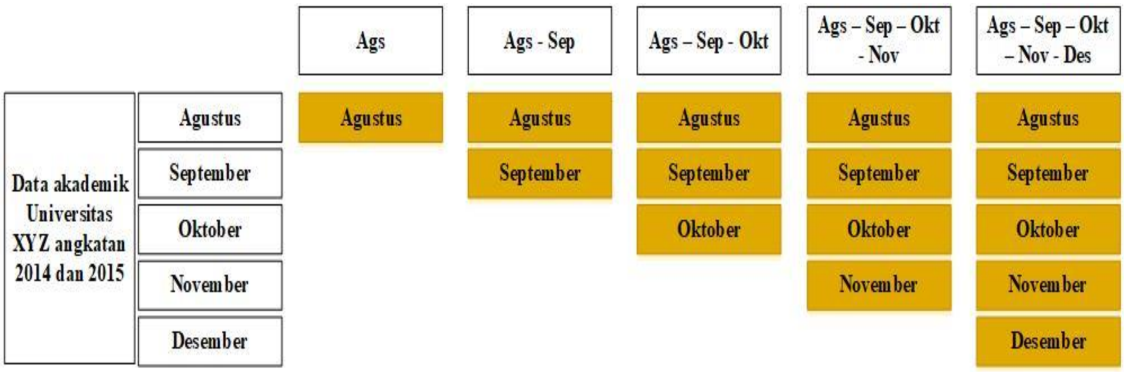


Gambar 3.2 Blok Diagram Penelitian

Data akademik akan dibagi menjadi data *training* dan *testing.* Data *training* diproses guna membentuk pola untuk validasi data *testing.* Data *training* diproses dengan metode *oversampling* MWMOTE untuk memisahkan kelas mayoritas dan minoritas (Sivaranjani, 2016). *Clustering* dilakukan dengan menggunakan *agglomerative hierarchical clustering* (AHC) *average*, *complete*, dan *single linkage*. Hasil *clustering* akan membentuk data sintetik dari *oversampling* yang berada dalam satu *cluster* (Barua *et al.*, 2014). Gabungan data minoritas hasil *clustering* (data sintetik) dan data minoritas akan memiliki rasio yang sama dengan data mayoritas. *Cophenetic correlation coefficient* merupakan metode yang digunakan dalam evaluasi AHC untuk mengetahui kinerja hasil *cluster*, sedangkan kinerja *oversampling* MWMOTE dievalusi dengan klasifikasi menggunakan *disscion tree* (J48) berdasarkan nilai *precision*, *recall*, *F-measure* dan *accuracy*.

3.2.1 *Dataset*

Dataset yang digunakan merupakan data akademik mahasiswa Universitas XYZ angkatan 2014 dan 2015. **Gambar 3.3** menjelaskan usulan data akademik yang digunakan dalam rancangan pemodelan deteksi periodik dengan atribut yang berbeda untuk setiap data.



Gambar 3.3 Skenario akumulasi data akademik Universitas XYZ Angkatan 2014 dan 2015

Tabel 3.1 Merupakan ilustrasi dataset yang akan digunakan dalam penelitian. Dataset akademik terdiri dari bulan Agustus, September, Oktober, November dan Desember, dimana data dijadikan data *trainning* dan *testing*. Model predeksi penanganan *imbalanced* menggunakan dataset setiap bulan untuk mengetahui performa dan kinerja metode *cluster* yang diusulkan pada algoritma MWMOTE. Agar data dapat diolah didalam penanganan *imbalanced*, maka perlu dilakukan beberapa preproses terhadap data mentah akademik. Data di proses dikarenakan bersumber dari 7 fakultas dan 24 program studi, sehingga terdapat perbedaan matakuliah dan aktivitas penunjang. Rata yang mempresentasikan keseluruhan data akademik

Tabel 3.1 Contoh Dataset Akademik Universitas XYZ Angkatan 2014 dan 2015

ID Mahasiswa	Absensi	Aktivitas Internet	Aktifitas Blog	Aktivitas Lomba & Panitia	Nilai UTS	Nilai Akhir	Label	Simbol
1	94.41	0.08	0	0	100	100	Lulus	1
2	11.48	0.18	0	0	77.70	61.11	Tidak Lulus	0
3	97.9	0.36	0	0	100	100	Lulus	1
...	…	…	…	…	…	…	…	
…	…	…	…	…	…	…	…	

3.2.2 *Pra-proses data imbalanced*

Pra-proses data *imbalanced* bertujuan mempersiapkan data akademik untuk diproses. Proses tersebut terdiri dari pemisahan data mayoritas dan data minoritas berdasarkan label kelas dataset (**Tabel 3.1**). Pra-proses kelas mayoritas berlabel lulus disimbolkan dengan angka satu (1), sedangkan angka nol (0) menunjukkan simbol kelas minoritas berlabel tidak lulus.

3.2.3 *Oversampling MWMOTE*

Oversampling MWMOTE melibatkan 3 fase, yaitu (Barua *et al.*, 2014):

Fase 1 Pemilihan sampel kelas minoritas.

1. S-min = Sampel data minoritas, S-maj = Sampel data mayoritas.
2. S-minf = Sampel data minoritas yang berada pada kelas mayoritas.

$$S_{minf} = S_{min} - \{x_i \in S_{min}: N\ N\ (x_i)\} \tag{3.1}$$

3. S-bmaj untuk menentukan *borderline* (batas) kelas mayor.

$$S_{bmaj} = \cup_{x_i \in S_{minf}} N_{maj}(x_i) \tag{3.2}$$

4. S-imin = Sampel data kelas minoritas informatife.

$$N_{min}(y_i)s\ S_{imin} = \cup_{y_i \in S_{bmaj}} N_{min}(y_i) \tag{3.3}$$

Fase 2 Pembobotan sampel kelas minoritas.

1. I-w = Pembobototan sampel data S-imin
2. S-w = Penyeleksian sampel data I-w sebagai dasar pembuatan data sintetik

$$S_w(x_i) = \sum_{y_i \in S_{bmaj}} I_w(y_i x_i) \tag{3.4}$$

3. S-p = Probabilitas sampel data S-w.

$$S_p(x_i) = \frac{S_w(x_i)}{\sum_{Z_i \in S_{imin}} S_w(Z_i)} \tag{3.5}$$

Fase 3 Pembuatan data sintetik sampel kelas minoritas menggunakan metode *clustering*

Pada penelitian ini, fase 3 digunakan untuk mengevaluasi *hierarchical clustering* guna memperoleh data sintetik yang lebih representatife pada sampel kelas minoritas.

1. Pilih data minoritas “x” dari S_{imin} sesuai dengan distribusi probabilitas $\{S_p(x_i)\}$, notasi “x” adalah anggota dari *cluster* C_k, $1 \leq k \leq$ M
2. Pilih data minoritas “y” secara random dari anggota *cluster* C_k

3.3 Skenario Uji Coba dan Pengujian

Pengujian berfungsi menangani *imbalanced* dataset akademik (**Tabel 3.1**) dan mengevaluasi *hierarchical clustering* pada MWMOTE.

1. *Imbalanced* dataset akademik ditangani dengan pembentukan data sintetik berdasarkan sampel kelas minoritas yang diproses dengan metode MWMOTE.
2. Pada AHC, seluruh sampel kelas minoritas merupakan *cluster* awal. *Cluster* awal tersebut nantinya akan diproses secara hiraki sampai terbentuk *cluster* tunggal.
3. *Clustering* yang digunakan AHC berupa *average*, *complete*, dan *single linkage*. Contoh proses *clustering* dengan AHC adalah sebagai berikut :
   a. *Average linkage clustering* sampel data minoritas pada data akademik

$$d_{C_{i,j}} = \frac{1}{|C_i| \times |C_j|} \sum_{i \in C_i, j \in C_j} d_{ij} \qquad (3.6)$$

Keterangan :

$C_{i,j}$ = *cluster*

d_{ij} = jarak anggota antar *cluster*

i, j = anggota *cluster*

Pada tahap awal seluruh sampel data minoritas merupakan *cluster* awal. Sampel data minoritas 3 dan 4 memiliki jarak terdekat yaitu 81.2109 sehingga dapat dijadikan dalam satu *cluster* (**Gambar 3.4**). Proses *clustering* dilanjutkan sampai diperoleh *cluster* dengan jumlah tertentu berdasarkan ambang batas minimum setiap anggota *cluster*.

$$
\begin{array}{cc}
 & \begin{array}{cccccc} 1 & 2 & 3 & 4 & 5 & 6 \end{array} \\
\begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{array} &
\begin{bmatrix}
0 & & & & & \\
85.2102 & 0 & & & & \\
589.5407 & 571.7396 & 0 & & & \\
587.9841 & 578.1380 & 81.2109 & 0 & & \\
720.6073 & 710.1504 & 157.4909 & 133.0670 & 0 & \\
861.8442 & 850.0266 & 293.1995 & 280.8847 & 155.1774 & 0
\end{bmatrix}
\end{array}
$$

Gambar 3.4 Contoh matrik data minoritas

Berikut merupakan perhitungan *average linkage* data minoritas 3 dan 4.

$$d_{3,4} = \frac{\{d_3 + \ d_4\}}{2} = \frac{\{81.2109 + 81.2109\}}{2} = 81.2109$$

b. *Complete linkage clustering* sampel data minoritas (**Gambar 3.4**) berfungsi membentuk *cluster* berbasis jarak terjauh antar sampel data kelas minoritas (**Persamaan 3.7**).

$$d_{C_{i,j}} = \ max_{i \in C_i, j \in C_j} d_{ij} \tag{3.7}$$

Berikut perhitungan *complete linkage* jarak terjauh data minoritas 1 dan 6 :

$$d_{1,6} = \max\{d_1, d_6\} = \max\{861.8442, 861.8442\} = 861.8442$$

c. *Single linkage clustering*, membentuk *cluster* dengan ketentuan jarak terdekat antar sampel data kelas minoritas (**Persamaan 3.8**).

$$d_{C_{i,j}} = \ min_{i \in C_i, j \in C_j} d_{ij} \tag{3.8}$$

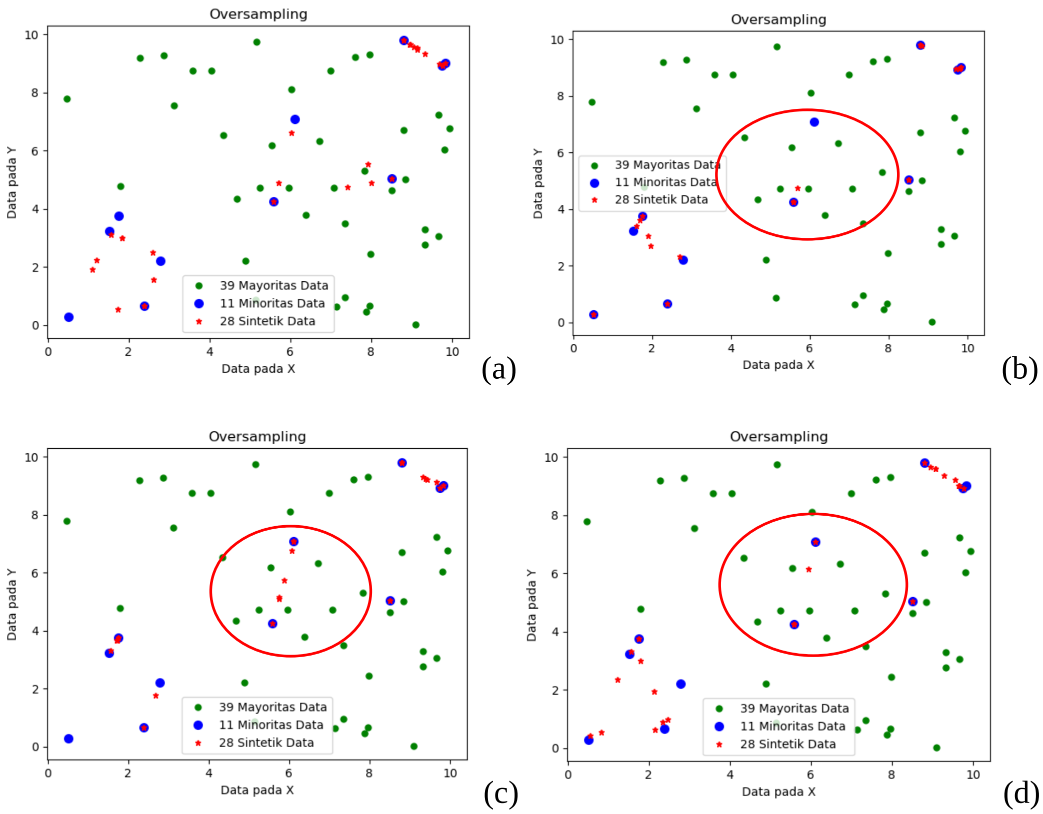
Data minoritas 3 dan 4 memiliki jarak data kelas minoritas terdekat (**Gambar 3.4**).

$$d_{3,4} = \min\{d_3, d_4\} = \min\{81.2109, 81.2109\} = 81.2109$$

Penggunaan ilustrasi dataset sintetik dengan mengkondisikan keseluruhan data, dimana data minoritas, mayoritas dan sintetik ditentukan dari awal (Tabel 3.2). Data tersebut merupakan dataset yang sudah dikondisikan terlebih dahulu dengan simbol yang telah ditentukan (Gambar 3.5(a)). Data minoritas disimbolkan dengan nol (0), data mayoritas dengan simbol satu (1), dan tanda nol bintang (0*) untuk data sintetik. Data tersebut berfungsi untuk memvalidasi hasil pembuatan data sintetik dengan algoritma MWMOTE.

Tabel 3.2 Ilustrasi Data Sintetik

Instance	Data X	Data Y	Label	Instance	Data X	Data Y	Label	Instance	Data X	Data Y	Label
1	0.51	0.28	0	16	7.36	3.50	1	31	1.10	1.91	0*
2	7.36	0.96	1	17	9.67	7.23	1	32	9.06	9.57	0*
3	7.15	0.63	1	18	0.27	9.20	1	33	8.81	9.80	0*
4	7.87	0.46	1	19	1.80	4.78	1	34	8.81	9.80	0*
5	5.17	9.74	1	20	6.02	8.11	1	35	9.69	8.98	0*
6	7.96	0.67	1	21	5.15	0.87	1	36	1.83	2.99	0*
7	1.76	3.76	0	22	3.12	0.10	0	37	5.59	4.25	0*
8	1.52	3.24	0	23	9.67	3.05	1	38	6.03	6.61	0*
…	…	…	…	…	…	…	…	…	…	…	…
…	…	…	…	…	…	…	…	…	…	…	…
…	…	…	…	…	…	…	…	…	…	…	…



Gambar 3.5 Data Sintetik (a), Data Sintetik dengan *Average linkage* (b),Data Sintetik dengan *Complete linkage* (c), dan Data Sintetik dengan *Single linkage* (d).

Pembuatan data sintetik untuk average linkage masih berada pada data mayor (**Gambar 3.5**(b) Lingkaran merah). Data sintetik pada complete linkage sama dengan data minoritas dan berada di data mayoritas (**Gambar 3.5**(c) Lingkaran merah). Pada single linkage data sintetik terbentuk berada di data mayoritas (**Gambar 3.5**(d) Lingkaran merah).

4. *Cophenetic correlation coefficient* (Persamaan 3.4) digunakan untuk pengujian *cluster* yang terbentuk dari data minoritas.

$$CCC = \frac{\sum_{i<j}(d_{ij} - \overline{d_{ij}})(d_{c_{ij}} - \overline{d_{c_{ij}}})}{\sqrt{\left[\sum_{i<j}(d_{ij} - \bar{d})^2\right]\left[\sum_{i<j}(d_{c_{ij}} - \overline{d_{c_{ij}}})^2\right]}} \tag{3.9}$$

CCC = *cophenetic correlation coefficient*

d_{ij} = jarak antar anggota i dan j

$d_{c_{ij}}$ = jarak dendogram yang terbentuk antar anggota i dan j

$\overline{d_{ij}}$ = rata-rata d_{ij}

$\overline{d_{c_{ij}}}$ = rata-rata $d_{c_{ij}}$

Berikut contoh perhitungan *cophenetic correlation coefficient.*

Tabel 3.3 Contoh data minoritas (Data akademik) Universitas XYZ

	Cluster							
	1	2	3	4	5	…	…	…
Dij	35.1794	48.0495	70.8995	85.6702	100.8263	…	…	…
dcij	35.1794	48.04945	63.2249	70.8995	108.1155	…	…	…

$$\bar{d} = \frac{1}{5}\,(35.1794 + 48.0495 + 70.8995 + 85.6702 + 100.8263) = 68.1250$$

$$\overline{d_c} = \frac{1}{5}\,(35.1794 + 48.0495 + 63.2249 + 70.8995 + 108.1155) = 65.0938$$

$$CCC = \frac{\sum_{i<j}(100.8263 - 68.1250)(108.115 - 65.0938)}{\sqrt{\left[\sum_{i<j}(100.8263 - 68.125)^2\right]\left[\sum_{i<j}(108.1155 - 65.0938)^2\right]}}$$

$$CCC = \frac{2831.265}{\sqrt{8831086.89}} = \frac{2831.265}{2971.714} = 0.9527$$

Cophenetic memiliki nilai $-1 \leq CCC \leq 1$. Nilai *cophenetic* mendekati 1 maka *cluster* disebut baik dan nilai *cophenetic* mendekati 0 maka *cluster* yang dihasilkan buruk (Bouguettaya *et al.*, 2015). Dari **Tabel 3.3** didapat nilai *cophenetic* 0.9527 menunjukkan bahwa *cluster* yang terbentuk adalah *cluster* optimum.

5. Data sintetik terbentuk dari pembobotan dan *clustering* menggunakan *Agglomerative Hierarchical Clustering* (AHC) kelas minoritas pada MWMOTE. Data sintetik yang terbentuk merupakan selisih antara data mayoritas dengan data minoritas (**Persamaan 3.10**).

$$for\ s\ in\ xrange(S_{maj} - S_{\min})$$

$$s_i = x_i + \alpha \times (y_i - x_i) \tag{3.10}$$

$$end$$

Keterangan :

s = Data sintetik

x = Sampel data kelas minoritas

α = Nilai random $0 \leq \alpha \leq 1$

y = Sampel data kelas minoritas dari *cluster* yang sama dengan x

3.4 Evaluasi Kinerja Metode

Evaluasi kinerja metode yang di usulkan pada data akademik tahun angkatan 2014 dan 2015 menggunakan *confusion matrix* **Tabel 3.4**.

Tabel 3.4 *Confusion Matrix*

Actual	(+) Predicted	(-) Predicted
(+)	TP	FN
(-)	FP	TN

Kelas (+) ditujukan untuk kelas mahasiswa yang tidak mengalami gagal akademik, yaitu mahasiswa lulus tahun pertama secara normal lebih banyak dari pada tidak lulus. Sedangkan kelas (-) merupakan kelas mahasiswa pada tahun pertama tidak lulus dengan permasalahan perkuliahan. Evaluasi kinerja yang digunakan *precision, recall, F-Measure,* dan *accuracy* untuk masing-masing kelas. *True positive* (TP) adalah kasus di mana prediksi ya (memiliki penyakit), dan memiliki penyakit. *True negative* (TN) memperkirakan tidak, dan tidak memiliki penyakit. *False positive* (FP) memperkirakan ya, tetapi sebenarnya tidak memiliki penyakit. (Juga dikenal sebagai "Tipe I kesalahan.") *False Negative* (FN) memperkirakan tidak, tetapi benar-benar memiliki penyakit. (Juga dikenal sebagai "Kesalahan Tipe II.")

$$Precision = \frac{TP}{(TP+FP)} \tag{3.11}$$

$$Recall = \frac{TP}{(TP+TN+FP+FN)} \tag{3.12}$$

$$F - Measure = 2 \times \frac{Precision.Recall}{(Precision+Recall)} \tag{3.13}$$

$$Accuracy = \frac{(TP+TN)}{(TP+TN+FP+FN)} \tag{3.14}$$

Precision adalah tingkat ketepatan antara informasi yang diminta oleh pengguna dengan jawaban yang diberikan oleh sistem. *Recall* adalah tingkat keberhasilan sistem dalam menemukan kembali sebuah informasi. F-*measure* merupakan salah satu perhitungan evaluasi dalam temu kembali informasi yang mengkombinasikan *recall* dan *precision*. *Accuracy* didefinisikan sebagai tingkat kedekatan antara nilai prediksi dengan nilai aktual.

BAB 4
HASIL PENELITIAN DAN PEMBAHASAN

Pada bagian ini dipaparkan hasil penelitian dan pembahasan dari setiap langkah yang telah dipaparkan pada **BAB 3**. Hasil penelitian dievaluasi sesuai dengan skenario pengujian yakni perhitungan *cophenetic correlation coefficient*, *precision*, *recall,* F-*measure*, dan *accuracy*.

4.1 Implementasi

Metode yang diusulkan diimplementasikan menggunakan bahasa pemrograman python 2.7 dan IDE Anaconda Navigator serta Sublime Text 3. Spesifikasi perangkat keras terdiri dari system operasi Windows 10 Enterprise 64-bit, RAM 16 GB, Processor Intel(R) Core(TM) i5 – 7400 CPU 3.00GHz.

4.2 Dataset Pengujian

Uji coba dilakukan pada data akademik angkatan 2014 dan 2015 Universitas XYZ sebanyak 12854 baris data dan 10 dataset UCI (**Tabel 4.1 dan 4.2**). Data tersebut merupakan akumulasi data akademik yang bersifat *imbalanced* pada bulan Agustus, September, Oktober, November, dan Desember. *Oversampling* dilakukan guna menangani *imbalanced* dataset. Pra-proses dilakukan sebelum *oversampling* MWMOTE dengan tujuan pemisahan data menjadi data *trainning* dan data *testing* dengan perbandingan persentase 70% : 30%. Data *trainning* digunakan untuk pembuatan model terhadap data *testing*. Data akademik memiliki dua label yaitu lulus dan tidak lulus, dalam tahap pra-proses label akan dirubah ke dalam angka 1 (lulus) dan 0 (tidak lulus) (**Lampiran 1**).

Tabel 4.1 Deskripsi Data akademik angkatan 2014 dan 2015 Universitas XYZ

<table>
<tr><th>Dataset</th><th>Atribut</th><th>Jumlah Data</th><th>Mayoritas</th><th>Minoritas</th><th>Persentase Minoritas</th><th>Imbalanced ratio</th></tr>
<tr><td>Agustus</td><td rowspan="2">34</td><td rowspan="5">12854</td><td rowspan="5">9482</td><td rowspan="5">3372</td><td rowspan="5">26%</td><td rowspan="5">0.74 : 0.26</td></tr>
<tr><td>September</td></tr>
<tr><td>Oktober</td><td rowspan="2">37</td></tr>
<tr><td>November</td></tr>
<tr><td>Desember</td><td>42</td></tr>
</table>

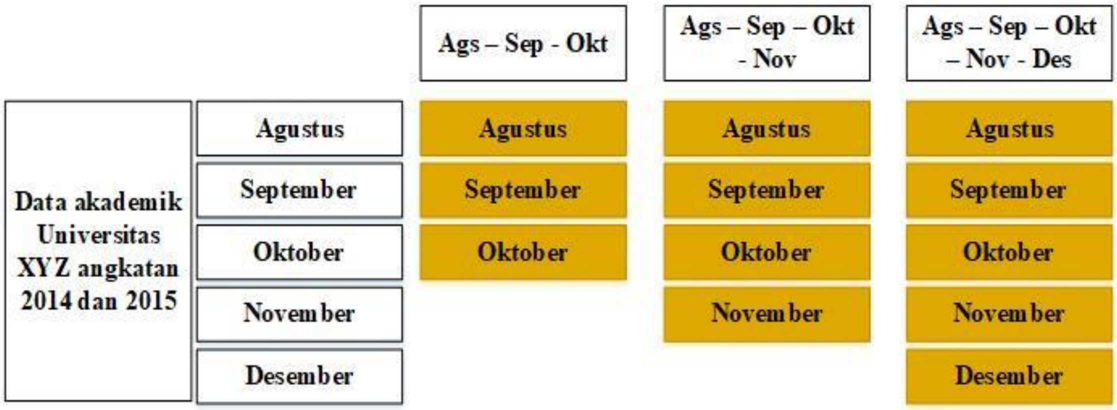
Tabel 4.2 Diskripsi 10 Dataset UCI

Dataset	Atribut	Jumlah Data	*Imbalanced* Ratio
Abalone	8	731	0.94 : 0.06
Breast	10	106	0.66 : 0.34
Ecoli	8	336	0.77 : 0.23
Glass	10	214	0.76 : 0.24
Libra	91	360	0.80 : 0.20
OCR	65	3823	0.90 : 0.10
Robot	25	5456	0.78 : 0.22
Satimage	37	6435	0.68 : 0.32
Wine	14	178	0.76 : 0.24
Yeast	9	1484	0.79 : 0.21

Dataset pengujian disusun ke dalam bentuk akumulasi bulanan (**Gambar 3.3**). Proses pencarian pola mahasiswa bermasalah dengan menggunakan klasifikasi data mining dilakukan terhadap data bulanan yang disusun terakumulasi dari satu bulan ke bulan-bulan berikutnya. Adanya perbedaan atribut bulan Agustus sampai dengan Desember dikarenakan ada pembagian data akademik. Pembagian 3 data berdasarkan bulan awal, bulan tengah semester dan akhir semester (**Gambar 4.1** dan **Lampiran 1**). Skenario tengah semester melengkapi skenario data akumulasi bulanan. Hasil nilai UTS dimasukkan ke dalam skenario bulanan akumulasi pada bulan Oktober, November, dan Desember (**Gambar 4.2**).

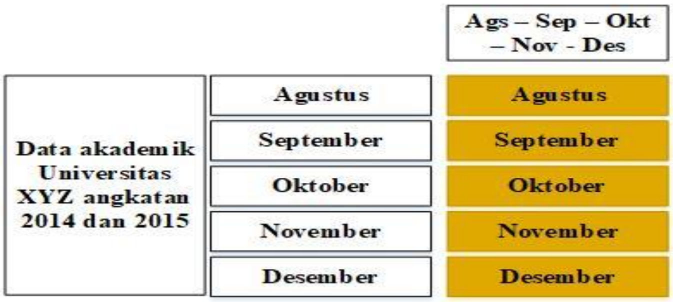
Bulan Awal	Bulan Tengah Semester	Bulan Akhir Semester
• Presensi • Aktivitas Internet • Aktivitas Blog • Aktivitas Lomba dan Panitia Media Sosial	• Presensi • Aktivitas Internet • Aktivitas Blog • Aktivitas Lomba dan Panitia Media Sosial • Nilai UTS	• Presensi Akhir • Aktivitas Internet • Aktivitas Blog • Aktivitas Lomba dan Panitia Media Sosial • Nilai Akhir • Transkrip Aktivitas Kemahasiswaan

Gambar 4.1 Atribut Data Akademik



Gambar 4.2 Akumulasi Bulan Tengah Semester

Skenario akhir semester melengkapi skenario bulanan dan tengah semester pada bulan Desember berupa tambahan nilai akhir dalam bentuk persentase Satuan Kredit Semester (SKS) lulus di akhir semester. Selain itu, digunakan pula persentase nilai di atas IPK 2, dan persentase nilai di bawah IPK 2 (**Gambar 4.3**)



Gambar 4.3 Akumulasi Bulan Akhir Semester

4.3 *Oversampling* MWMOTE

Oversampling MWMOTE memiliki tiga fase dalam menangani *imbalanced* dataset. Dalam penanganan *imbalanced* data akademik dan dataset UCI dilakukan iterasi sebanyak 15 kali untuk mengetahui data sintetik yang terbentuk memiliki hasil representatif (**Lampiran 2 dan 4**).

1. Fase pertama

   Pada fase ini, dilakukan pemisahan antara dataset minoritas / S_min (**Tabel 4.2**) dengan dataset mayoritas / S_maj (**Tabel 4.3**). Data minoritas dan mayoritas dipisahkan dengan adanya 2 label kelas yaitu lulus dan tidak lulus. Label kelas tidak lulus memiliki jumlah data yang lebih sedikit yang dianggap sebagai data minoritas dan sebaliknya untuk mahasiswa dengan label lulus sebagai data mayoritas (**Lampiran 1**). Data minoritas dibagi menjadi tiga bagian dengan ketentuan berada pada data mayoritas, dekat dengan *borderline*, dan data minoritas yang berada pada *borderline*.

Tabel 4.3 Contoh Dataset Minoritas

Index Data	Atribut 1	Atribut 2	Atribut 3	Atribut 4	…	…	Atribut 'n'
2	11.48	17.43	53.8	75.83	…	…	39.635
4	82.52	38.45	52.93	86.51	…	…	61.11
22	18.03	17.43	53.8	75.83	…	…	36.352
31	43.51	72.19	69.74	77.06	…	…	11.11
32	45.38	31.77	57.27	75.83	…	…	44.272
33	96.75	72.19	69.74	77.06	…	…	61.11
34	89.92	32.08	52.17	84.06	…	…	66.67
35	96.12	32.08	52.17	84.06	…	…	100
49	95.24	28.26	53.8	75.48	…	…	83.33
62	93.23	18.75	61.71	65.85	…	…	66.67
…	…	…	…	…	…	…	…
…	…	…	…	…	…	…	…

Pada pre-proses, data minoritas memiliki 3 kriteria, yaitu data minoritas yang berada dalam data mayoritas (S_minf), data minoritas yang berada pada *borderline* (S_bmaj), dan data yang bersifat informatif (S_imin).

Tabel 4.4 Contoh Dataset Mayoritas

Index Data	Atribut 1	Atribut 2	Atribut 3	Atribut 4	…	…	Atribut 'n'
0	94.41	53.42	62.91	78.28	…	…	100
1	98.59	51.79	62.8	78.28	…	…	100
3	93.65	28.26	53.8	75.48	…	…	100
5	98.59	51.79	62.8	78.28	…	…	100
6	100	51.79	62.8	78.28	…	…	100
7	96.38	37.71	75.92	63.05	…	…	100
8	100	32.21	52.71	82.14	…	…	100
9	99.18	17.43	53.8	75.83	…	…	100
16	98.59	75.92	77.81	80.21	…	…	90
18	93.65	75.48	65.9	91.23	…	…	90
20	98.59	78.28	88.97	78.32	…	…	90
21	100	57.67	66.75	93.26	…	…	100
…	…	…	…	…	…	…	…
…	…	…	…	…	…	…	…
…	…	…	…	…	…	…	…

2. Fase kedua pembobotan data minoritas

Pembobotan data minoritas digunakan dalam penentuan kandidat data sintetik (**Tabel 4.5**). Pembobotan memiliki 3 bagian :

a. Pembobotan data minoritas yang mempunyai jarak terdekat dengan *borderline.*
b. Kelompok data minoritas terdiri dari kelompok kecil.
c. Kelompok data minoritas berjumlah banyak berada dalam data mayoritas.

Tabel 4.5 Contoh Pembobotan Dataset Minoritas

No	Anggota 1 (Mayoritas)	Anggota 2 (Minoritas)	Hasil Pembobotan	No	Anggota 1 (Mayoritas)	Anggota 2 (Minoritas)	Hasil Pembobotan
1	25	16	0.001	16	12	21	0.11087
2	27	21	0.05895	17	25	29	0.01656
3	13	4	0.00934	18	27	19	0.0014
4	18	4	0.00742	19	15	4	0.00924
5	3	21	0.07148	20	9	16	0.00095
6	9	29	0.01404	21	15	29	0.01109
7	18	19	0.00149	22	27	4	0.00844
8	9	4	0.00785	23	3	19	0.0013
…	…	…	…	…	…	…	…
…	…	…	…	…	…	…	…
…	…	…	…	…	…	…	…
…	…	…	…	…	…	…	…
…	…	…	…	…	…	…	…

3. Fase ketiga proses *clustering*

MWMOTE menggunakan proses *clustering* untuk membentuk kumpulan data minoritas dalam satu kelompok yang memiliki kemiripan data. Metode *cluster* yang digunakan adalah sebagai berikut :

a. *Average linkage*, menentukan anggota *cluster* dengan menggunakan jarak antar data berdasarkan rata-rata (**Tabel 4.6**).
b. *Complete linkage clustering*, menentukan anggota *cluster* berdasarkan jarak terjauh antara data (**Tabel 4.7**).
c. Penentuan anggota cluster berdasarkan jarak terdekat antar data disebut *single linkage* (**Tabel 4.8)**.

Tabel 4.6 Proses *Clustering* dengan *Average Linkage*

Proses	Hasil
Cluster Awal	0: [2], 1: [4], 2: [22], 3: [31], 4: [32], 5: [33], 6: [34], 7: [35], 8: [49], 9: [62], 10: [77], 11: [81], 12: [82], 13: [83], 14: [84], 15: [85], 16: [86], 17: [87], 18: [91], 19: [104], 20: [112], 21: [113], 22: [114], 23: [115], 24: [116], 25: [117], 26: [118], …, …, …, …
Gabung Cluster	2: 0
Gabung Cluster	104: 0, 2: 0
Gabung Cluster	104: 0, 2: 0, 22: 0
Gabung Cluster	104: 0, 2: 0, 126: 0, 22: 0
Gabung Cluster	104: 0, 32: 0, 2: 0, 126: 0, 22: 0
Gabung Cluster	32: 0, 2: 0, 22: 0, 104: 0, 86: 0, 126: 0
Gabung Cluster	32: 0, 2: 0, 4: 1, 22: 0, 145: 1, 104: 0, 34: 1, 81: 1
Gabung Cluster	32: 0, 2: 0, 4: 1, 22: 0, 145: 1, 104: 0, 34: 1, 81: 1, 117: 0
…	…
…	…
Jumlah Cluster	312
Anggota cluster	129: 1, 2: 0, 4: 1, 133: 1, 135: 1, 138: 1, 139: 1, 140: 1, 141: 1, 145: 1,, …, …, …, …

Tabel 4.7 Proses *Clustering* dengan *Complete Linkage*

Proses	Hasil
Cluster Awal	0: [2], 1: [4], 2: [22], 3: [31], 4: [32], 5: [33], 6: [34], 7: [35], 8: [49], 9: [62], 10: [77], 11: [81], 12: [82], 13: [83], 14: [84], 15: [85], 16: [86], 17: [87], 18: [91], 19: [104], 20: [112], 21: [113], 22: [114], 23: [115], 24: [116], 25: [117], 26: [118], …, …, …, …
Gabung Cluster	2: 0
Gabung Cluster	104: 0, 117: 0
Gabung Cluster	104: 0, 117: 0, 86: 0
Gabung Cluster	104: 0, 117: 0, 86: 0, 118: 1
Gabung Cluster	104: 0, 34: 1, 117: 0, 86: 0, 118: 1
Gabung Cluster	104: 0, 34: 1, 81: 1, 117: 0, 86: 0, 118: 1
…	…
…	…
…	…
Jumlah Cluster	270
Anggota cluster	135: 1, 138: 1, 139: 1, 140: 1, 141: 1, 145: 1, 22: 0, 31: 3, 32: 0, 33: 1, …, …, …, …

Tabel 4.8 Proses *Clustering* dengan *Single Linkage*

Proses	Hasil
Cluster Awal	0: [2], 1: [4], 2: [22], 3: [31], 4: [32], 5: [33], 6: [34], 7: [35], 8: [49], 9: [62], 10: [77], 11: [81], 12: [82], 13: [83], 14: [84], 15: [85], 16: [86], 17: [87], 18: [91], 19: [104], 20: [112], 21: [113], 22: [114], 23: [115], 24: [116], 25: [117], 26: [118], …, …, …, …
Gabung Cluster	22: 0
Gabung Cluster	22: 0, 104: 0
Gabung Cluster	22: 0, 145: 1, 104: 0
Gabung Cluster	22: 0, 145: 1, 104: 0, 34: 1
Gabung Cluster	22: 0, 32: 0, 33: 1, 34: 1, 35: 1
Gabung Cluster	22: 0, 32: 0, 33: 1, 34: 1, 35: 1, 49: 1
…	…
…	…
…	…
Jumlah Cluster	219
Anggota cluster	87: 1, 91: 1, 104: 0, 112: 20, 113: 1, 114: 14, 115: 1, 116: 1, 117: 0, …, …, …, …

Clustering pada MWMOTE menggunakan *bottom-up* yang dimulai dari seluruh data minoritas sebagai *cluster*. Selanjutnya menggabungkan menjadi satu *cluster* besar dengan menentukan *threshold* (T_h) untuk menjadikan batas ambang jumlah *cluster* yang terbentuk. Penggunaan *threshold* bertujuan menghasilkan sejumlah *cluster* yang berbeda untuk jenis data yang sama, di mana satu-satunya perbedaan adalah dimensi fitur. *Threshold* yang konstan akan mendapatkan fakkta bahwa dalam beberapa dataset relatif jarang (jarak rata-rata antara sampel yang besar), sementara di dataset lainnya relatif padat (rata-rata dari sampel yang berbeda-beda). *Threshold* bergantung pada data dan harus dihitung menggunakan beberapa heuristik dari ukuran jarak antara data. Nilai *threshold* didapat dari jarak rata-rata minimum (S_{minf}) dari seluruh data (d_{avg}) dan nilai konstanta (C_p). C_p merupakan nilai konstan yang diberikan untuk menentukan jumlah *cluster* yang terbentuk. Semakin tinggi nilai C_p maka penurunan terhadap jumlah *cluster*, akan tetapi bertambah jumlah anggota *cluster* dan skenario sebaliknya akan terjadi jika nilai C_p.

$$d_{avg} = \frac{1}{|S_{minf}|} \sum_{x \in S_{minf}} \min_{y \neq x, y \in S_{minf}} \{dist(x, y)\} \tag{4.1}$$

$$T_h = d_{avg} \times C_p \tag{4.2}$$

Tabel 4.9 merupakan kandidat pembuatan data sintetik berdasarkan *cluster* dan pembobotan (S_w) pada tahap kedua MWMOTE secara random dalam satu *cluster*. Kandidat data sintetik memiliki bobot tertentu untuk menyatakan data sintetik akan dibuat. Data sintetik terbentuk dari gabungan anggota *cluster* dari data minoritas (**Tabel 4.10 dan 4.11**). Jumlah data sintetik yang terbentuk merupakan jumlah selisih antara data mayoritas dengan data minoritas. Hasil *oversampling* memiliki rasio yang seimbang dengan perbandingan data mayoritas dengan data minoritas serta data sintetik.

Tabel 4.9 Data Sintetik berdasarkan Pembobotan (S_w)

No.	Kandidat Cluster	Anggota Cluster	Pemilihan Bobot (S_w)
1	133	1	0.010421395329481827
2	135	1	0.06226312077758183
3	138	1	0.03734055088768434
…	…	…	…
…	…	…	…
…	…	…	…

Tabel 4.10 Contoh Pembuatan Data Sintetik

Anggota cluster =	129: 1, 2: 0, 4: 1, 133: 1, 135: 1, 138: 1, 139: 1, 140: 1, 141: 1, 145: 1, 22: 0, 31: 3, 32: 0, 33: 1, 34: 1, 35: 1, 49: 1, 62: 1, 77: 1, 81: 1, …, …, …, …, …)
Pembuatan Data Sintetik ke 1	
Anggota Cluster – 1	116
Anggota Cluster -2	138
Alpha	0.1
Data Sintetik	[95.029, 59.609, 67.99, 68.58500000000001, 77.78, 62.777]
Pembuatan Data Sintetik ke 2	
Anggota Cluster – 1	135
Anggota Cluster -2	83
Alpha	0.51
Data Sintetik	[93.9318, 35.738299999999995, 62.298, 71.94069999999999, 91.8317, 91.8317]
…	
…	…
…	…
…	…
…	…

Tabel 4.11 Contoh Data Sintetik yang Terbentuk

	Atribut 1	Atribut 2	Atribut 3	Atribut 4	Atribut 5	Atribut 6
Sintetik 1	89.0072	33.2531	53.3737	80.8847	89.1122	72.4422
Sintetik 2	82.52	38.45	52.93	86.51	77.78	61.11
Sintetik 3	26.5568	19.3794	53.8	75.767	18	14.9994
Sintetik 4	22.1075	17.6016	54.8283	74.5326	10.8329	8.6671
Sintetik 5	82.5168	31.3456	56.4712	72.4392	53.3352	42.8868
Sintetik 6	94.2551	23.6001	57.6759	70.7613	91.8317	75.1666
Sintetik 7	84.5552	36.8196	53.0692	84.7452	81.3352	64.6652
Sintetik 8	67.7723	29.8005	53.3749	78.0774	61	50.8313
Sintetik 9	74.792	33.219	58.093	70.593	25.003	18.332
Sintetik 10	48.914	21.762	53.8	75.69	50.002	33.332
…	…	…	…	…	…	…
…	…	…	…	…	…	…

4.4 Evaluasi Ratio MWMOTE

Berdasarkan data akademik, terdapat 5 data yang terdiri dari bulan Agustus, September, Oktober, November, dan Desember. Data akademik memiliki rasio yang sama dari ke 5 data tersebut sebesar 0.74 : 0.26, dimana 0.74 merupakan data mayoritas dan 0.26 adalah data minoritas (**Tabel 4.1**). Data akademik yang memiliki karakteristik data yang berbeda antara Agustus – Desember, sehingga data mayoritas dan minoritas memiliki hasil evaluasi berbeda. Data UCI yang digunakan adalah Abalone, Breast, Ecoli, Glass, Libra, OCR, Robot, Satimage, Wine dan Yeast dimana data UCI tersebut memiliki jumlah rasio yang tidak seimbang antara kelas mayoritas dan minoritas (**Tabel 4.2**). Untuk data UCI terlihat berbeda mulai dari jumlah atribut, jumlah data, dan *imbalanced Ratio* akan menyebabkan permasalahan dalam *imbalanced*. Hasil pengujian metode klasifikasi di dataset yang tidak seimbang biasanya memiliki ciri khas berupa nilai *instance* yang terklasifikasi (*misclassification cost*) di kelas minoritas lebih tinggi dibandingkan dengan *misclassification cost* di kelas mayoritas. Diperlukan penangaan data *imbalanced* untuk mendapatkan hasil klasifikasi yang representatif dan akurat.

MWMOTE dapat digunakan dalam penangaan data *imbalanced* dengan cara pembuatan data sintetik. Data sintetik terbentuk dari proses oversampling berdasarkan pembobotan dan pengklasteran dari data minoritas. Sehingga didapat data akademik dan UCI yang memiliki rasio seimbang berdasarkan jumlah data (**Tabel 4.12** dan **Tabel 4.13**).

Selisih antara data mayoritas dengan data minoritas dijadikan acuan untuk pembuatan data sintetik.

Tabel 4.12 Rasio Jumlah Data Akademik Tanpa MWMOTE dan Dengan MWMOTE

Imbalaced data	*Imbalanced* ratio Tanpa MWMOTE		*Imbalanced* ratio dengan MWMOTE		
	Mayoritas	Minoritas	Mayoritas	Minoritas	Sintetik
Data Akademik Agustus	0.74	0.26	0.74	0.26	0.52
Data Akademik September					
Data Akademik Oktober					
Data Akademik November					
Data Akademik Desember					

Tabel 4.13 Rasio Jumlah Data UCI Tanpa MWMOTE dan Dengan MWMOTE

Imbalaced data	*Imbalanced* ratio Tanpa MWMOTE		*Imbalanced* ratio dengan MWMOTE		
	Mayoritas	Minoritas	Mayoritas	Minoritas	Sintetik
Data UCI Abalone	0.94	0.06	0.94	0.06	0.88
Data UCI Breast	0.66	0.34	0.66	0.34	0.32
Data UCI Ecoli	0.77	0.23	0.77	0.23	0.54
Data UCI Glass	0.76	0.24	0.76	0.24	0.52
Data UCI Libra	0.80	0.20	0.80	0.20	0.60
Data UCI OCR	0.90	0.10	0.90	0.10	0.80
Data UCI Robot	0.78	0.22	0.78	0.22	0.56
Data UCI Satimage	0.68	0.32	0.68	0.32	0.36
Data UCI Wine	0.76	0.24	0.76	0.24	0.52
Data UCI Yeast	0.79	0.21	0.79	0.21	0.58

Data yang sudah seimbang akan diproses dengan menggunakan *classifier dicission tree* (J48). Untuk mengetahuai algoritma MWMOTE dapat menangani permasalahan *imbalanced* pada data akademik dan data UCI.

4.5 Evaluasi dan Pengujian

Evaluasi dan pengujian metode pada penelitian ini digunakan untuk mengetahui apakah metode yang diusulan tersebut dapat menyelesaikan permasalahan *imbalanced* pada kasus data akademik Univeristas XZY angkatan 2014 dan 2015 dan dataset UCI. Evaluasi terbagi menjadi 2 macam yaitu evaluasi *cluster* dan evaluasi algoritma MWMOTE. Pada penelitian ini didapatkan hasil evaluasi yang berbeda antara *cluster* dan metode dengan *classifier*.

4.5.1 *Evaluasi Cluster*

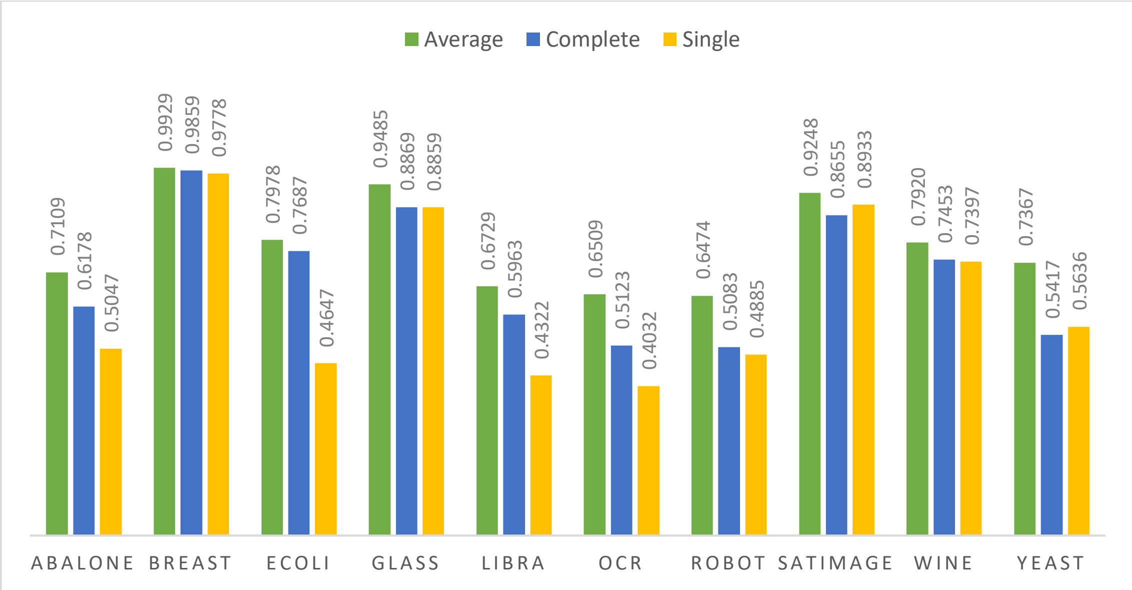
Evaluasi *cluster* dilakukan dengan *Cophenetic Correlation Coefficient* (CCC). Berdasarkan hasil uji coba, digunakan metode *hierarchical clustering* untuk mengetahui performa algoritma MWMOTE guna mendapatkan data sintetik yang representative berdasarkan hasil akurasi dari klasifikasi *dicission tree*. Data sintetik terbentuk dari data minoritas dan data pada *cluster* (kemiripan data dalam satu *cluster*). **Tabel 4.14** menunjukkan *average linkage clustering* menghasilkan nilai *Cophenetic Correlation Coefficient* yang lebih besar dibandingkan dengan *complete* dan *single linkage*.

Hasil evaluasi *cluster* dilakukan berdasarkan data minoritas yang memiliki kesamaan antar anggota *cluster* dengan cara menganalisis seluruh data. Penentuan CCC antar pasangan data *cluster* dilakukan melalui pengukuran antar anggota *cluster.* Nilai CCC terbaik pada data akademik adalah 0.94350 dengan *average*, 0.90986 dengan *complete* dan 0.89377 dengan *single* di bulan agustus. Nilai CCC dikatakan baik apabila mendekati 1.

Tabel 4.14 Hasil Evaluasi *Cluster* dengan CCC

	AGUSTUS	SEPTEMBER	OKTOBER	NOVEMBER	DESEMBER
Average	**0.94350**	**0.83730**	**0.84525**	**0.86386**	**0.82656**
Complete	0.90986	0.56457	0.72691	0.62255	0.69469
Single	0.89377	0.65250	0.75661	0.80126	0.67243

Evaluasi metode *cluster* dilakukan lebih lanjut dengan uji coba menggunakan UCI *Machine Learning Repository* dataset. UCI *Machine Learning Repository* adalah dataset umum yang banyak dipakai para peneliti untuk mengevaluasi metode yang diusulkan. Pada penelitian ini, dataset UCI digunakan untuk validasi hasil analisis *cluster.* Abalone, Breast, Ecoli, Glass, Libra, OCR, Robot, Satimage, Wine dan Yeast merupakan dataset UCI yang dipakai untuk mengetahui kinerja *cluster*. **Gambar 4.4** menunjukkan nilai tertinggi CCC terdapat pada *average linkage*. Hasil evaluasi data dengan menggunakan dataset UCI menunjukkan bahwa *average linkage* memiliki nilai tertinggi dari *complete* dan *single linkage* pada data minoritas. Guna mengetahui kesamaan hasil data sintetik yang terbentuk dari data minoritas hasil CCC, dilakukan evaluasi *classification* dengan J48 pada data *imbalanced*.



Gambar 4.4 Hasil evaluasi CCC dengan Dataset UCI

4.5.2 *Evaluasi Metode*

Evaluasi metode yang diusulkan menggunakan *decision tree classification* (J48) untuk data akademik Universitas XYZ angkatan 2014 dan 2015. Pengujian evaluasi meliputi beberapa kriteria, yaitu nilai *precision, recall, F-Measure*, dan *accuracy*. **Tabel 4.12** menyajikan data akademik dengan rasio yang tidak seimbang. Dilakukan evaluasi untuk mengetahui perbaik dari data yang tidak seimbang. Berdasarkan pengujian menunjukkan penggunaan *single linkage* pada algoritma MWMOTE memiliki nilai paling tinggi dibandingkan *average* dan *complete linkage*. Berdasarkan data akademik per bulan, evaluasi *imbalanced* dataset akademik bulan Desember dengan algoritma MWMOTE menghasilkan nilai tertinggi berdasarkan seluruh kriteria dibandingkan dengan bulan yang lainnya (menggunakan *average*, *complete*, dan *single linkage*) (**Tabel 4.15 - 4.17**).

Tabel 4.15 Data Akademik sebelum dilakukan *Oversampling*

	Precision	*Recall*	*F-measure*	*Accuracy*
Desember	72.00%	70.70%	72.03%	72.00%
November	71.20%	70.90%	71.21%	71.20%
Oktober	71.10%	70.40%	71.13%	71.10%
September	70.80%	70.00%	70.83%	70.80%
Agustus	70.00%	70.20%	70.10%	70.16%

Tabel 4.16 Hasil Evaluasi *Imbalanced* dataset dengan *Average Linkage* dan MWMOTE

	Precision	*Recall*	*F-measure*	*Accuracy*
Desember	**88.26%**	**87.94%**	**88.08%**	**87.95%**
November	79.39%	78.14%	78.61%	78.14%
Oktober	78.63%	77.53%	77.98%	77.53%
September	73.25%	72.89%	73.03%	72.88%
Agustus	70.37%	71.93%	70.91%	71.93%

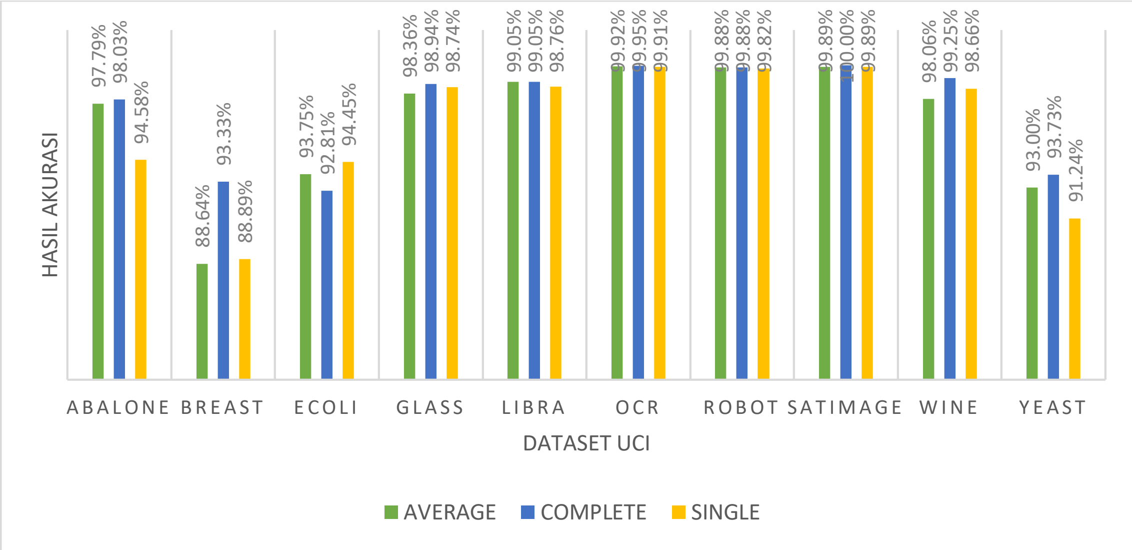
Tabel 4.17 Hasil Evaluasi *Imbalanced* dataset dengan *Complete Linkage* dan MWMOTE

	Precision	*Recall*	*F-measure*	*Accuracy*
Desember	**88.24%**	**87.89%**	**88.01%**	**87.88%**
November	79.53%	78.17%	78.69%	78.17%
Oktober	78.68%	77.67%	78.07%	77.66%
September	73.37%	73.13%	73.25%	73.13%
Agustus	70.27%	71.27%	70.67%	71.28%

Tabel 4.18 Hasil Evaluasi *Imbalanced* dataset dengan *Single Linkage* dan MWMOTE

	Precision	*Recall*	*F-measure*	*Accuracy*
Desember	**88.19%**	**87.81%**	**87.95%**	**87.81%**
November	79.48%	78.22%	78.69%	78.22%
Oktober	78.87%	77.85%	78.24%	77.84%
September	73.39%	73.38%	73.38%	73.37%
Agustus	70.17%	71.49%	70.70%	71.50%

Performa metode yang diusulkan divalidasi dengan data UCI *Machine Learning Repository* (Abalone, Breast, Ecoli, Glass, Libra, OCR, Robot, Satimage, Wine dan Yeast). **Gambar 4.5** adalah hasil evaluasi akurasi pada 10 dataset UCI. Evaluasi dari dataset UCI didapat hasil yang berbeda dikarenakan setiap dataset memiliki karakteristik berbeda. Pada *complete linkage* memiliki hasil evaluasi akurasi tertinggi dari 7 dataset yaitu abalone, breast, glass, ocr, satimage, wine, dan yeast. Hasil *average linkage* pada data UCI memiliki rata-rata 96.83% dari seluruh dataset. Akurasi rata-rata dari *single linkage* untuk seluruh data UCI adalah 96.49%.



Gambar 4.5 Hasil Evaluasi Akurasi pada Dataset UCI

4.6 Uji Statistik

Uji statistika menjadi uji pembanding untuk menentukan nilai perbedaan signifikan antar *average*, *complete*, dan *single linkage* pada *oversampling* MWMOTE. Pada penelitian ini, pengujian statistika dilakukan pada salah satu kriteria yaitu *accuracy*. Metode *cluster* yang diusulkan dapat meperbaiki hasil evaluasi pada algoritma MWMOTE merupakan H_0. Jika tidak perbedaan signifikan pada uji statistik dinyatakan H_1. Hipotesis (H_0 dan H_1) dapat dinyatakan berdasarkan nilai *alpha* (p-*value/sig*). Jika H_0 diterima nilai *alpha* < 0.05 dan sebaliknya H_0 ditolak jika nilai *alpha* (p-*value/sig*) > 0.05. Uji statistika diawali dengan uji normalitas dan homogenitas sebagai syarat pengujian *One Way* ANOVA (Sopiyudin, 2001)(**Lampiran 2**).

Data dinyatakan normal dan homogen apabila memiliki nilai *alpha* (p-*value/sig*) > 0.05. Hasil *One Way* ANOVA dapat diuji lanjut dengan *Duncan* apabila memiliki nilai *alpha* (p-*value/sig*) < 0.05. Berdasarkan hasil *One Way* ANOVA, metode *average*, *complete*, dan *single linkage* tidak signifikan terhadap peningkatan nilai *accuracy* pada algoritma *oversampling* MWMOTE. Hasil tersebut didapatkan berdasarkan keseluruhan data akademik pada bulan Agustus (**Tabel 4.19**) dengan nilai p-*value* 0.178, September (**Tabel 4.20**) dengan nilai p-*value* 0.211, Oktober (**Tabel 4.21**) dengan nilai p-*value* 0.484, November (**Tabel 4.22**) dengan nilai p-*value* 0.946, dan Desember (**Tabel 4.23**) dengan nilai p-*value* 0.749.

Pada data UCI dengan pengujian *One Way* ANOVA didapat hasil bahwa Abalone, Robot, Satimage memiliki nilai sig/*alpha* < 0.05. Pada pengujian lanjut dapat diteruskan dengan uji lanjut Duncan berdasarkan nilai sig/*alpha* untuk menyatakan perbedaan signifikan terjadi pada *single* linkage terhadap average dan *complete linkage* (**Lampiran 5**). Breast, Ecoli, Glass, Libra, OCR, Wine dan Yeast tidak memiliki beda signifikan berdasarkan uji ANOVA dengan sig/*alpha* > 0.05 (**Lampiran 5**),

Tabel 4.19 Hasil Uji *One Way* ANOVA Bulan Agustus dengan Nilai Sig > *0.05*

	Sum of Squares	df	*Mean Square*	F	Sig.
Between Groups	.000	2	.000	1.796	**.178**
Within Groups	.004	42	.000		
Total	.004	44			

Tabel 4.20 Hasil *Uji One Way* ANOVA Bulan September dengan Nilai Sig > 0.05

	Sum of Squares	df	*Mean Square*	F	Sig.
Between Groups	.000	2	.000	1.617	**.211**
Within Groups	.002	42	.000		
Total	.002	44			

Tabel 4.21 Hasil Uji *One Way* ANOVA Bulan Oktober dengan Nilai Sig > 0.05

	Sum of Squares	df	*Mean Square*	F	Sig.
Between Groups	.000	2	.000	.738	**.484**
Within Groups	.002	42	.000		
Total	.002	44			

Tabel 4.22 Hasil Uji *One Way* ANOVA Bulan November dengan Nilai Sig > 0.05

	Sum of Squares	df	*Mean Square*	F	Sig.
Between Groups	.000	2	.000	.055	**.946**
Within Groups	.002	42	.000		
Total	.002	44			

Tabel 4.23 Hasil Uji *One Way* ANOVA Bulan Desember dengan Nilai Sig > 0.05

	Sum of Squares	df	*Mean Square*	F	Sig.
Between Groups	.000	2	.000	.291	**.749**
Within Groups	.001	42	.000		
Total	.001	44			

BAB 5
KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil uji coba dari metode usulan memberikan beberapa kesimpulan yaitu :

1. Metode *oversampling* MWMOTE digunakan untuk menangani permasalahan *imbalanced* data akademik Universitas XYZ angkatan 2014 dan 2015 dengan nilai akurasi 87.95% pada *average*, 87.88% pada *complete* dan *single* sebesar 87.81%. Data UCI dalam penelitian diperoleh hasil evaluasi rata-rata akurasi 96.83% untuk *average,* 97.50% pada *complete,* dan 96.49% *single*.
2. Algoritma MWMOTE terbukti memperbaiki akurasi rata-rata sebesar 6.61% dari seluruh bulan pada data akademik dan data UCI sebesar 1.2% dibandingkan tanpa penanganan *imbalanced*.
3. *Average, complete* dan *single* dapat digunakan pada MWMOTE dengan tingkat akurasi yang tidak berbeda signifikan berdasarkan uji statistic *One Way* ANOVA dengan nilai *sign* / *alpha* > 0.005
4. Evaluasi *cluster* terbaik didapat pada metoda *average linkage* dengan nilai *cophenetic* sebesar 0.8632 pada data akademik dan 0.78747 pada data UCI.

5.2 Saran

Berdasarkan hasil yang diperoleh pada penelitian ini, perlu dilakukan pengkajian ulang dan penggunaan kombinasi metode dalam memperbaiki algoritma MWMOTE sehingga didapatkan nilai yang sama pada evalusi *cluster* ataupun evaluasi metode *oversampling.*

DAFTAR PUSTAKA

1. Abd Rahman, N., Abu Bakar, Z. and Zulkefli, N. S. S. (2016) 'Malay document clustering using complete linkage clustering technique with Cosine Coefficient', *ICOS 2015 - 2015 IEEE Conference on Open Systems*, pp. 103–107. doi: 10.1109/ICOS.2015.7377286.
2. Almeida, J. A. S. *et al.* (2007) 'Improving hierarchical cluster analysis: A new method with outlier detection and automatic clustering', *Chemometrics and Intelligent Laboratory Systems*, 87(2), pp. 208–217. doi: 10.1016/j.chemolab.2007.01.005.
3. Barua, S. *et al.* (2014) 'MWMOTE - Majority weighted minority oversampling technique for imbalanced data set learning', *IEEE Transactions on Knowledge and Data Engineering*, 26(2), pp. 405–425. doi: 10.1109/TKDE.2012.232.
4. Beyan, C. and Fisher, R. (2015) 'Classifying imbalanced data sets using similarity based hierarchical decomposition', *Pattern Recognition*. Elsevier, 48(5), pp. 1653–1672. doi: 10.1016/j.patcog.2014.10.032.
5. Blagus, R. *et al.* (2013) 'SMOTE for high-dimensional class-imbalanced data', *BMC Bioinformatics*, 14(1), p. 106. doi: 10.1186/1471-2105-14-106.
6. Bouguettaya, A. *et al.* (2015) 'Efficient agglomerative hierarchical clustering', *Expert Systems with Applications*, 42(5), pp. 2785–2797. doi: 10.1016/j.eswa.2014.09.054.
7. Chawla, N. V. *et al.* (2002) 'SMOTE: Synthetic minority over-sampling technique', *Journal of Artificial Intelligence Research*, 16, pp. 321–357. doi: 10.1613/jair.953.
8. El-Sayed, A. A. *et al.* (2016) 'Handling autism imbalanced data using synthetic minority over-sampling technique (SMOTE)', *Proceedings of 2015 IEEE World Conference on Complex Systems, WCCS 2015*, pp. 1–5. doi: 10.1109/ICoCS.2015.7483267.
9. Fachrurrozi, M. *et al.* (2017) 'The grouping of facial images using agglomerative hierarchical clustering to improve the CBIR based face recognition system', *2017 International Conference on Data and Software Engineering (ICoDSE)*, pp. 1–6. doi: 10.1109/ICODSE.2017.8285868.
10. Fahrudin, T. (2017) 'DETEKSI DINI KEGAGALAN AKADEMIK SERTA MULTILABELISASI PERMASALAHAN MAHASISWA DARI DATA MEDIA SOSIAL'.
11. Fahrudin, T., Buliali, J. L. and Fatichah, C. (2016) 'RANDSHUFF : an Algorithm to Handle Imbalance Class for Qualitative Data', 11(December), pp. 1093–1104. doi: 10.15866/irecos.v11i12.10956.
12. Fahrudin, T., Buliali, J. L. and Fatichah, C. (2017) 'Predictive modeling of the first year evaluation based on demographics data: Case study students of Telkom University, Indonesia', *Proceedings of 2016 International Conference on Data and Software Engineering, ICoDSE 2016*, pp. 0–5. doi: 10.1109/ICODSE.2016.7936158.
13. Fakhruzi, I. (2018) 'An Artificial Neural Network with Bagging to Address Imbalance Datasets on Clinical Prediction', *2018 International Conference on Information and Communications Technology (ICOIACT)*, (1), pp. 895–898. doi: 10.1109/ICOIACT.2018.8350824.
14. Gong, J. and Kim, H. (2017) 'RHSBoost: Improving classification performance in imbalance data', *Computational Statistics and Data Analysis*. Elsevier B.V., 111, pp. 1–13. doi: 10.1016/j.csda.2017.01.005.

15. Guo, S. *et al.* (2016) ‘A centroid-based gene selection method for microarray data classification’, *Journal of Theoretical Biology*. Elsevier, 400, pp. 32–41. doi: 10.1016/j.jtbi.2016.03.034.
16. Hakim, R. B. F., Subanar and Winarko, E. (2010) ‘Reducing hierarchical clustering instability using clustering based on indiscernibility and indiscernibility level’, *Proceedings - 2010 IEEE International Conference on Granular Computing, GrC 2010*, pp. 182–187. doi: 10.1109/GrC.2010.136.
17. Huth, R., Nemesova, I. and Klimperova, N. (1993) ‘Weather Categorization Based on the Average Linkage Clustering Technique - an Application to European Midlatitudes’, *International Journal of Climatology*, 13(8), pp. 817–835.
18. Irawan, E. *et al.* (2015) ‘Penggunaan Random Under Sampling untuk Penanganan Ketidakseimbangan Kelas pada Prediksi Cacat Software Berbasis Neural Network’, *Journal of Software Engineering*, 1(2), pp. 92–100.
19. Jayasree, S. and Gavya, A. A. (2015) ‘Classification of Imbalance Problem by MWMOTE and SSO’, pp. 1–4.
20. John, M. and Jayasudha, J. S. (2017) ‘Enhancing Performance of Deep Learning Based Text Summarizer’, *International Journal of Applied Engineering Research*, 12(24), pp. 15986–15993.
21. Köhn, H.-F. and Hubert, L. J. (2015) ‘Hierarchical Cluster Analysis’, *Wiley StatsRef: Statistics Reference Online*, pp. 1–13. doi: 10.1002/9781118445112.stat02449.pub2.
22. Kumar, S. and Toshniwal, D. (2016) ‘Analysis of hourly road accident counts using hierarchical clustering and cophenetic correlation coefficient (CPCC)’, *Journal of Big Data*. Springer International Publishing, 3(1), pp. 1–11. doi: 10.1186/s40537-016-0046-3.
23. Larose, D. T. and Larose, C. D. (2014) ‘Discovering Knowledge in data An Introduction to Data Mining’, *Journal of Chemical Information and Modeling*, 53(9), pp. 1689–1699. doi: 10.1017/CBO9781107415324.004.
24. Lesmana, I. P. D. (2012) ‘Pengembangan Decision Tree J48 Untuk Diagnosis Penyakit Diabetes Mellitus’, *Seminar Nasional Teknologi Informasi & Komunikasi Terapan*, 2012(Semantik), pp. 189–193.
25. Li, K., Wang, L. and Hao, L. (2009) ‘Comparison of cluster ensembles methods based on hierarchical clustering’, *Proceedings of the 2009 International Conference on Computational Intelligence and Natural Computing, CINC 2009*, (1), pp. 499–502. doi: 10.1109/CINC.2009.214.
26. Lin, T. W. *et al.* (2017) ‘Differential Covariance: A New Class of Methods to Estimate Sparse Connectivity from Neural Recordings’, 2394, pp. 2379–2394. doi: 10.1162/NECO.
27. Liu W, Chawla, N. V. and C. D. (2010) ‘A Robust Decision Tree Algorithm for Imbalanced Data Sets’, *Society for Industrial and Applied Mathematics*, pp. 766–777. doi: 10.1137/1.9781611972801.67.
28. Ma, L. and Fan, S. (2017) ‘CURE-SMOTE algorithm and hybrid algorithm for feature selection and parameter optimization based on random forests’, *BMC Bioinformatics*, 18(1), pp. 1–18. doi: 10.1186/s12859-017-1578-z.
29. Mahmood, A. M. (2015) ‘Class Imbalance Learning in Data Mining – A Survey’, *International Journal of Communication Technology for Social Networking Services*, 3(2), pp. 17–36. doi: 10.21742/ijctsns.2015.3.2.02.
30. Mellor, A. *et al.* (2015) ‘Exploring issues of training data imbalance and mislabelling on random forest performance for large area land cover classification using the

ensemble margin', *ISPRS Journal of Photogrammetry and Remote Sensing*. International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS), 105, pp. 155–168. doi: 10.1016/j.isprsjprs.2015.03.014.

31. Napierała, K. (2012) 'Improving Rule Classifiers For Imbalanced Data', (October).
32. Ng, W. W. Y. *et al.* (2015) 'Diversified sensitivity-based undersampling for imbalance classification problems', *IEEE Transactions on Cybernetics*, 45(11), pp. 2402–2412. doi: 10.1109/TCYB.2014.2372060.
33. Piri, S., Delen, D. and Liu, T. (2018) 'A synthetic informative minority over-sampling (SIMO) algorithm leveraging support vector machine to enhance learning from imbalanced datasets', *Decision Support Systems*. Elsevier B.V, 106, pp. 15–29. doi: 10.1016/j.dss.2017.11.006.
34. Purwar, A. and Singh, S. K. (2015) 'Hybrid Prediction Model with missing value Imputation for medical data', *EXPERT SYSTEMS WITH APPLICATIONS*. Elsevier Ltd. doi: 10.1016/j.eswa.2015.02.050.
35. Pusadan, M Y, Buliali, J L, Ginardi, and R. V. H. (2016) 'ANOMALY DETECTION OF FLIGHT ROUTES THROUGH OPTIMAL WAYPOINT', *International Conference on Computing and Applied Informatics 2016*, 9(1), pp. 3–10. doi: 10.1088/1742-6596/801/1/012041.
36. Ristoski, P. and Paulheim, H. (2016) 'Semantic Web in data mining and knowledge discovery: A comprehensive survey', *Journal of Web Semantics*. Elsevier B.V., 36, pp. 1–22. doi: 10.1016/j.websem.2016.01.001.
37. Rivera, W. A. (2017) 'Noise Reduction A Priori Synthetic Over-Sampling for class imbalanced data sets', *Information Sciences journal*. Elsevier Inc., 408, pp. 146–161. doi: 10.1016/j.ins.2017.04.046.
38. Seiffert, C., Khoshgoftaar, T. M. and Van Hulse, J. (2009) 'Hybrid sampling for imbalanced data', *Integrated Computer-Aided Engineering*, 16(3), pp. 193–210. doi: 10.3233/ICA-2009-0314.
39. Sivaranjani, S. (2016) 'GIS based Serial Crime Analysis using Data Mining Techniques', 153(8), pp. 19–23.
40. Sopiyudin, D. (2001a) 'Statistik untuk Kedokteran dan Kesehatan Edisi Tiga'. Jakarta, Indonesia: Penerbit Salemba Medika, pp. 1–47. Available at: https://www.bukusopiyudin.com.
41. Zheng, Z., Cai, Y. and Li, Y. (2015) 'Oversampling method for imbalanced classification', *Computing and Informatics*, 34(5), pp. 1017–1037.

Lampiran 1

1. Data akademik Universitas XYZ Angkatan 2014 – 2015 bulan Agustus

Instance	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12	A13	A14	A15	A16	A17	A18	A19	A20	A21	A22	A23
1178	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1184	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1190	0.0	0.0	0.0	0.0	0.0	66.7	0.0	0.0	33.3	0.0	0.0	0.0	50.0	50.0	0.0	0.0	0.0	0.0	50.0	50.0	0.0	0.0	0.0
1192	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1196	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1199	0.0	0.0	0.0	0.0	0.0	69.2	0.0	0.0	30.8	0.0	0.0	0.0	25.0	0.0	25.0	50.0	50.0	25.0	25.0	0.0	0.0	0.0	0.0
1203	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1206	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1210	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1211	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1213	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1215	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1225	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1226	0.0	0.0	0.0	0.0	0.0	87.5	0.0	0.0	12.5	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0
1228	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1229	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1230	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1232	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1233	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1236	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1240	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1241	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1245	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1251	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Instance	A24	A25	A26	A27	A28	A29	A30	A31	A32	A33	A34	Label
1178	0.0	0.0	0.0	0.0	0.0	0.0	0.0	43.5	58.2	100.0	0	Tidak Lulus
1184	0.0	0.0	0.1	0.0	0.5	0.0	0.0	60.9	58.2	100.0	0	Tidak Lulus
1190	0.0	0.0	0.0	0.0	0.1	0.0	0.0	14.9	34.4	80.0	0	Tidak Lulus
1192	0.0	0.0	0.0	0.0	0.0	0.0	0.0	10.4	46.7	80.0	1	Lulus
1196	0.0	0.0	0.0	0.0	0.0	0.0	0.0	15.2	41.8	100.0	1	Lulus
1199	0.0	0.0	0.0	0.0	0.0	0.0	0.0	70.9	67.4	100.0	1	Lulus
1203	0.0	0.0	0.1	0.0	0.3	0.0	0.2	60.9	58.2	100.0	0	Tidak Lulus
1206	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0	Tidak Lulus
1210	0.0	0.0	0.1	0.0	0.3	0.2	0.0	25.2	46.1	80.0	1	Lulus
1211	0.0	0.0	0.8	0.0	3.2	0.0	0.0	43.5	58.2	100.0	0	Tidak Lulus
1213	0.0	0.0	0.0	0.0	0.0	0.0	0.0	11.1	58.7	80.0	0	Tidak Lulus
1215	0.0	0.0	0.0	0.0	0.0	0.0	0.0	24.9	57.4	60.0	0	Tidak Lulus
1225	0.0	0.0	0.0	0.0	0.0	0.0	0.0	15.2	36.3	80.0	1	Lulus
1226	0.0	0.0	0.0	0.0	0.0	0.0	0.0	15.2	41.8	100.0	1	Lulus
1228	0.0	0.0	0.0	0.0	0.0	0.0	0.0	25.2	46.1	80.0	1	Lulus
1229	0.0	0.0	0.1	0.0	0.4	0.0	0.0	14.9	34.4	80.0	1	Lulus
1230	0.0	0.0	0.0	0.0	0.0	0.0	0.0	43.9	63.3	80.0	1	Lulus
1232	0.0	0.0	0.0	0.0	0.0	0.0	0.0	60.9	58.2	100.0	1	Lulus
1233	0.0	0.0	0.2	0.0	0.0	0.6	0.0	11.1	58.7	80.0	1	Lulus
1236	0.0	0.0	0.9	0.0	2.6	0.6	0.0	10.4	46.7	80.0	1	Lulus
1240	0.0	0.0	1.4	0.0	3.0	1.7	0.0	14.9	34.4	80.0	1	Lulus
1241	0.0	0.0	0.0	0.0	0.0	0.0	0.0	15.2	36.3	80.0	1	Lulus
1245	0.0	0.0	9.5	0.0	16.8	13.4	1.8	25.2	46.1	80.0	1	Lulus
1251	0.0	0.0	1.4	0.0	5.5	0.0	0.0	60.9	58.2	100.0	1	Lulus

2. Data akademik Universitas XYZ Angkatan 2014 – 2015 bulan September

Instance	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12	A13	A14	A15	A16	A17	A18	A19	A20	A21	A22	A23
1178	0.0	0.0	0.0	0.0	0.0	89.5	0.0	8.8	1.8	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0
1184	0.0	0.0	0.0	0.0	0.0	98.3	0.0	0.0	1.8	0.0	0.0	0.0	0.0	0.0	0.0	100.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0
1190	0.0	0.0	0.0	0.0	0.0	19.6	0.0	0.0	80.4	0.0	19.5	34.2	12.2	19.5	0.0	14.6	53.7	0.0	34.2	12.2	0.0	0.0	0.0
1192	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1196	0.0	0.0	0.0	0.0	0.0	96.0	0.0	0.0	4.0	0.0	50.0	0.0	0.0	50.0	0.0	0.0	50.0	0.0	50.0	0.0	0.0	0.0	0.0
1199	0.0	0.0	0.0	0.0	0.0	81.3	0.0	0.0	18.8	0.0	0.0	0.0	16.7	16.7	33.3	33.3	33.3	16.7	25.0	25.0	0.0	0.0	0.0
1203	0.0	0.0	0.0	0.0	0.0	98.3	0.0	0.0	1.8	0.0	0.0	0.0	0.0	0.0	0.0	100.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0
1206	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1210	0.0	0.0	0.0	0.0	0.0	98.1	0.0	0.0	1.9	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0
1211	0.0	0.0	0.0	0.0	0.0	98.3	0.0	0.0	1.8	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0
1213	0.0	0.0	0.0	0.0	0.0	98.3	0.0	0.0	1.8	0.0	100.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0
1215	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1225	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1226	0.0	0.0	0.0	0.0	0.0	98.0	0.0	0.0	2.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0
1228	0.0	0.0	0.0	0.0	0.0	90.6	3.8	0.0	5.7	0.0	0.0	66.7	33.3	0.0	0.0	0.0	33.3	33.3	0.0	33.3	0.0	0.0	0.0
1229	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1230	0.0	0.0	0.0	0.0	0.0	98.1	0.0	0.0	1.9	0.0	0.0	100.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0
1232	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1233	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1236	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1240	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1241	0.0	0.0	0.0	0.0	0.0	98.0	0.0	0.0	2.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0
1245	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1251	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Instance	A24	A25	A26	A27	A28	A29	A30	A31	A32	A33	A34	Label
1178	0.0	0	0.01	0	0	0.02	0	67.12	70.29	83.75	0	Tidak Lulus
1184	0.0	0	0.05	0	0.08	0	0	63.94	70.29	83.75	0	Tidak Lulus
1190	0.0	0	0.01	0	0.02	0	0	31.04	63.24	80.66	0	Tidak Lulus
1192	0.0	0	0	0	0	0	0	42.43	59.88	93.42	1	Lulus
1196	0.0	0	0.01	0	0.02	0	0	39.97	62.66	80.66	1	Lulus
1199	0.0	0	0	0	0	0	0	67.72	67.17	96.71	1	Lulus
1203	0.0	0	0.05	0	0.05	0	0.09	63.94	70.29	83.75	0	Tidak Lulus
1206	0.0	0	0	0	0	0	0	0	0	0	0	Tidak Lulus
1210	0.0	0	0.17	0	0.05	0.5	0	24.54	65.9	80.66	1	Lulus
1211	0.0	0	0.53	0	0.56	0.59	0	67.12	70.29	83.75	0	Tidak Lulus
1213	0.0	0	0	0	0	0	0	46.96	85.9	67.7	0	Tidak Lulus
1215	0.0	0	0	0	0	0	0	44.7	69.02	74.08	0	Tidak Lulus
1225	0.0	0	0	0	0	0	0	42.44	55.95	93.42	1	Lulus
1226	0.0	0	0	0	0	0	0	39.97	62.66	80.66	1	Lulus
1228	0.0	0	0	0	0	0	0	24.54	65.9	80.66	1	Lulus
1229	0.0	0	0.15	0	0.23	0	0	31.04	63.24	80.66	1	Lulus
1230	0.0	0	0	0	0	0	0	63.21	80.69	67.7	1	Lulus
1232	0.0	0	0	0	0	0	0	63.94	70.29	83.75	1	Lulus
1233	0.0	0	3.34	2.26	1.32	3.87	3.01	46.96	85.9	67.7	1	Lulus
1236	0.0	0	0.51	0	0.6	0.44	0	42.43	59.88	93.42	1	Lulus
1240	0.0	0	0.65	0	0.53	1.1	0	31.04	63.24	80.66	1	Lulus
1241	0.0	0	0	0	0	0	0	42.44	55.95	93.42	1	Lulus
1245	0.0	0	8.87	0	3.6	22.3	1.42	24.54	65.9	80.66	1	Lulus
1251	0.0	0	0.68	0	1.06	0	0	63.94	70.29	83.75	1	Lulus

3. Data akademik Universitas XYZ Angkatan 2014 – 2015 bulan Oktober

Instance	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12	A13	A14	A15	A16	A17	A18	A19	A20	A21	A22	A23
1178	0.0	0.0	0.0	0.0	0.0	93.0	0.0	5.8	1.2	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0
1184	0.0	0.0	0.0	0.0	0.0	98.8	0.0	0.0	1.2	0.0	0.0	0.0	0.0	0.0	0.0	100.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0
1190	0.0	0.0	0.0	0.0	0.0	16.7	0.0	0.0	83.3	0.0	18.3	30.0	11.7	23.3	0.0	16.7	53.3	0.0	33.3	13.3	0.0	0.0	0.0
1192	0.0	0.0	0.0	0.0	0.0	96.1	0.0	0.0	3.9	0.0	33.3	0.0	0.0	0.0	66.7	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0
1196	0.0	0.0	0.0	0.0	0.0	95.8	0.0	0.0	4.2	0.0	33.3	0.0	0.0	66.7	0.0	0.0	33.3	0.0	66.7	0.0	0.0	0.0	0.0
1199	0.0	0.0	0.0	0.0	0.0	80.9	0.0	0.0	19.1	0.0	11.8	0.0	11.8	29.4	23.5	23.5	35.3	29.4	17.7	17.7	0.0	0.0	0.0
1203	0.0	0.0	0.0	0.0	0.0	98.8	0.0	0.0	1.2	0.0	0.0	0.0	0.0	0.0	0.0	100.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0
1206	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1210	0.0	0.0	0.0	0.0	0.0	94.7	0.0	0.0	5.3	0.0	25.0	25.0	0.0	50.0	0.0	0.0	25.0	0.0	50.0	25.0	0.0	0.0	0.0
1211	0.0	0.0	0.0	0.0	0.0	98.8	0.0	0.0	1.2	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0
1213	0.0	0.0	0.0	0.0	0.0	98.8	0.0	0.0	1.2	0.0	100.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0
1215	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1225	0.0	0.0	0.0	0.0	0.0	98.7	0.0	0.0	1.3	0.0	100.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0
1226	0.0	0.0	0.0	0.0	0.0	98.6	0.0	0.0	1.4	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0
1228	0.0	0.0	0.0	0.0	0.0	88.2	2.6	0.0	9.2	0.0	14.3	42.9	14.3	28.6	0.0	0.0	42.9	28.6	14.3	14.3	0.0	0.0	0.0
1229	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1230	0.0	0.0	0.0	0.0	0.0	97.4	0.0	0.0	2.6	0.0	0.0	50.0	50.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0
1232	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1233	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1236	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1240	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1241	0.0	0.0	0.0	0.0	0.0	93.4	0.0	0.0	6.6	0.0	20.0	0.0	0.0	40.0	40.0	0.0	40.0	20.0	0.0	40.0	0.0	0.0	0.0
1245	0.0	0.0	0.0	0.0	0.0	98.7	0.0	0.0	1.3	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	100.0	0.0	0.0
1251	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Instance	A24	A25	A26	A27	A28	A29	A30	A31	A32	A33	A34	A35	A36	A37	Label
1178	0.0	0.0	0.0	0.0	0.0	0.0	0.0	73.7	69.1	79.3	50.0	16.7	33.3	0	Tidak Lulus
1184	0.0	0.0	0.0	0.0	0.1	0.0	0.0	71.0	68.7	79.3	66.7	0.0	33.3	0	Tidak Lulus
1190	0.0	0.0	0.0	0.0	0.0	0.0	0.0	28.5	62.4	73.4	0.0	66.7	33.3	0	Tidak Lulus
1116	0.0	0.0	0.0	0.0	0.0	0.0	0.0	42.7	58.1	78.6	50.0	33.3	16.7	1	Lulus
1051	0.0	0.0	0.0	0.0	0.0	0.0	0.0	37.2	58.7	71.4	66.7	0.0	33.3	1	Lulus
1052	0.0	0.0	0.0	0.0	0.0	0.0	0.0	63.8	60.2	83.9	27.8	55.6	16.7	1	Lulus
1203	0.0	0.0	0.0	0.0	0.1	0.0	0.0	71.0	68.7	79.3	50.0	16.7	33.3	0	Tidak Lulus
1206	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0	Tidak Lulus
1050	0.0	0.0	0.1	0.0	0.1	0.3	0.0	25.2	68.9	64.3	50.0	16.7	33.3	1	Lulus
1211	0.0	0.0	0.2	0.0	0.6	0.3	0.0	73.7	69.1	79.3	50.0	16.7	33.3	0	Tidak Lulus
1213	0.0	0.0	0.0	0.0	0.0	0.0	0.0	46.5	83.5	63.6	50.0	16.7	33.3	0	Tidak Lulus
1215	0.0	0.0	0.0	0.0	0.0	0.0	0.0	46.9	65.6	66.1	83.3	0.0	16.7	0	Tidak Lulus
1035	0.0	0.0	0.0	0.0	0.0	0.0	0.0	44.7	52.7	82.1	44.4	38.9	16.7	1	Lulus
1036	0.0	0.0	0.1	0.0	0.0	0.2	0.1	37.2	58.7	71.4	50.0	16.7	33.3	1	Lulus
1037	0.0	0.0	0.0	0.0	0.0	0.0	0.0	25.2	68.9	64.3	33.3	33.3	33.3	1	Lulus
1038	0.0	0.0	0.2	0.0	0.4	0.1	0.3	28.5	62.4	73.4	83.3	0.0	16.7	1	Lulus
1039	0.0	0.0	0.0	0.0	0.0	0.0	0.0	61.0	72.6	60.7	50.0	16.7	33.3	1	Lulus
1040	0.0	0.0	0.1	0.0	0.3	0.3	0.0	71.0	68.7	79.3	66.7	0.0	33.3	1	Lulus
1041	0.0	0.0	3.9	2.8	2.3	5.5	5.2	46.5	83.5	63.6	66.7	0.0	33.3	1	Lulus
1042	0.0	0.0	0.3	0.0	0.7	0.3	0.0	42.7	58.1	78.6	83.3	0.0	16.7	1	Lulus
1043	0.0	0.0	0.5	0.0	0.9	0.7	0.0	28.5	62.4	73.4	66.7	16.7	16.7	1	Lulus
1044	0.0	0.0	0.0	0.0	0.0	0.0	0.0	44.7	52.7	82.1	27.8	55.6	16.7	1	Lulus
1045	0.0	0.0	4.4	0.0	4.4	12.1	0.6	25.2	68.9	64.3	44.4	22.2	33.3	1	Lulus
1046	0.0	0.0	0.5	0.0	1.1	0.5	0.0	71.0	68.7	79.3	66.7	0.0	33.3	1	Lulus
1047		0.0	0.0	0.0	0.0	0.0	0.0	73.7	69.1	79.3	50.0	16.7	33.3	1	Lulus
1049		0.0	0.0	0.0	0.1	0.0	0.0	71.0	68.7	79.3	66.7	0.0	33.3	1	Lulus
1115		0.0	0.0	0.0	0.0	0.0	0.0	28.5	62.4	73.4	0.0	66.7	33.3	1	Lulus

4. Data akademik Universitas XYZ Angkatan 2014 – 2015 bulan November

Instance	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12	A13	A14	A15	A16	A17	A18	A19	A20	A21	A22	A23
1178	0.0	0.0	0.0	0.0	0.0	93.8	0.0	3.9	2.3	0.0	0.0	0.0	0.0	33.3	33.3	33.3	0.0	33.3	66.7	0.0	0.0	0.0	0.0
1184	0.0	0.0	0.0	0.0	0.0	98.4	0.0	0.0	1.6	0.0	0.0	0.0	0.0	0.0	0.0	100.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0
1190	0.0	0.0	0.0	0.0	0.0	13.1	0.0	0.0	86.9	0.0	19.4	31.2	11.8	22.6	0.0	15.1	53.8	0.0	34.4	11.8	0.0	0.0	0.0
1192	0.0	0.0	0.0	0.0	0.0	95.0	0.0	0.0	5.0	0.0	16.7	0.0	0.0	0.0	33.3	50.0	83.3	16.7	0.0	0.0	0.0	0.0	0.0
1196	0.0	0.0	0.0	0.0	0.0	92.9	0.0	0.0	7.1	0.0	25.0	0.0	25.0	50.0	0.0	0.0	50.0	0.0	50.0	0.0	0.0	0.0	0.0
1199	0.0	0.0	0.0	0.0	0.0	80.8	0.0	0.0	17.7	1.5	17.4	8.7	8.7	26.1	17.4	21.7	34.8	39.1	13.0	13.0	0.0	0.0	0.0
1203	0.0	0.0	0.0	0.0	0.0	98.4	0.0	0.0	1.6	0.0	0.0	0.0	0.0	0.0	50.0	50.0	50.0	0.0	0.0	50.0	0.0	0.0	0.0
1206	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1210	0.0	0.0	0.0	0.0	0.0	95.8	0.0	0.0	4.2	0.0	20.0	20.0	20.0	40.0	0.0	0.0	20.0	20.0	40.0	20.0	0.0	0.0	0.0
1211	0.0	0.0	0.0	0.0	0.0	98.5	0.0	0.0	1.6	0.0	0.0	0.0	50.0	0.0	0.0	50.0	50.0	50.0	0.0	0.0	0.0	0.0	0.0
1213	0.0	0.0	0.0	0.0	0.0	98.4	0.0	0.0	1.6	0.0	50.0	0.0	0.0	50.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0
1215	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1225	0.0	0.0	0.0	0.0	0.0	96.6	0.0	0.9	2.6	0.0	33.3	0.0	0.0	0.0	0.0	66.7	100.0	0.0	0.0	0.0	0.0	0.0	0.0
1226	0.0	0.0	0.0	0.0	0.0	98.2	0.0	0.0	1.8	0.0	0.0	0.0	0.0	50.0	0.0	50.0	50.0	0.0	50.0	0.0	0.0	0.0	0.0
1228	0.0	0.0	0.0	0.0	0.0	87.3	1.7	3.4	7.6	0.0	22.2	33.3	11.1	33.3	0.0	0.0	44.4	22.2	22.2	11.1	0.0	0.0	0.0
1229	0.0	0.0	0.0	0.0	0.0	99.1	0.0	0.0	0.9	0.0	0.0	100.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0
1230	0.0	0.0	0.0	0.0	0.0	96.4	0.0	0.0	3.6	0.0	0.0	25.0	50.0	0.0	0.0	25.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0
1232	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1233	0.0	0.0	0.0	0.0	0.0	96.8	0.0	2.4	0.8	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0
1236	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1240	0.0	0.0	0.0	0.0	0.0	99.1	0.0	0.0	0.9	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0
1241	0.0	0.0	0.0	0.0	0.0	89.7	0.0	0.0	10.3	0.0	33.3	8.3	0.0	25.0	16.7	16.7	50.0	33.3	0.0	16.7	0.0	0.0	0.0
1245	0.0	0.0	0.0	0.0	0.0	99.2	0.0	0.0	0.9	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	100.0	0.0	0.0
1251	0.0	0.0	0.0	0.0	0.0	97.6	0.0	0.8	1.6	0.0	0.0	0.0	0.0	0.0	50.0	50.0	0.0	50.0	0.0	50.0	0.0	0.0	0.0

Instance	A24	A25	A26	A27	A28	A29	A30	A31	A32	A33	A34	A35	A36	A37	Label
1178	0.0	0.0	0.1	0.0	0.2	0.0	0.0	63.2	69.5	80.9	50.0	16.7	33.3	0	Tidak Lulus
1184	0.0	0.0	0.2	0.0	0.4	0.3	0.1	59.8	69.1	80.9	66.7	0.0	33.3	0	Tidak Lulus
1190	0.0	0.0	0.0	0.0	0.0	0.0	0.0	20.1	59.7	76.7	0.0	66.7	33.3	0	Tidak Lulus
1192	0.0	0.0	0.1	0.0	0.2	0.1	0.0	36.7	57.7	85.1	50.0	33.3	16.7	1	Lulus
1196	0.0	0.0	0.2	0.0	0.2	0.5	0.0	34.8	59.6	77.6	66.7	0.0	33.3	1	Lulus
1199	0.0	0.0	0.1	0.0	0.2	0.3	0.0	46.9	58.7	88.8	27.8	55.6	16.7	1	Lulus
1203	0.0	0.0	0.2	0.0	0.3	0.4	0.0	59.8	69.1	80.9	50.0	16.7	33.3	0	Tidak Lulus
1206	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0	Tidak Lulus
1210	0.0	0.0	0.1	0.0	0.1	0.2	0.0	22.4	69.0	70.1	50.0	16.7	33.3	1	Lulus
1211	0.0	0.0	0.3	0.0	0.6	0.5	0.0	63.2	69.5	80.9	50.0	16.7	33.3	0	Tidak Lulus
1213	0.0	0.0	0.0	0.0	0.0	0.1	0.0	42.3	84.1	64.6	50.0	16.7	33.3	0	Tidak Lulus
1215	0.0	0.0	0.0	0.0	0.0	0.0	0.0	40.6	65.8	71.3	83.3	0.0	16.7	0	Tidak Lulus
1225	0.0	0.0	0.0	0.0	0.0	0.0	0.0	36.3	53.7	87.6	44.4	38.9	16.7	1	Lulus
1226	0.0	0.0	0.3	0.0	0.6	0.5	0.2	34.8	59.6	77.6	50.0	16.7	33.3	1	Lulus
1228	0.0	0.0	0.0	0.0	0.1	0.0	0.0	22.4	69.0	70.1	33.3	33.3	33.3	1	Lulus
1229	0.0	0.0	0.3	0.0	0.6	0.3	0.2	20.1	59.7	76.7	83.3	0.0	16.7	1	Lulus
1230	0.0	0.0	0.8	0.0	1.1	1.8	0.3	44.8	71.9	61.3	50.0	16.7	33.3	1	Lulus
1232	0.0	0.0	0.1	0.0	0.3	0.2	0.0	59.8	69.1	80.9	66.7	0.0	33.3	1	Lulus
1233	0.0	0.0	7.1	8.7	3.6	6.9	9.4	42.3	84.1	64.6	66.7	0.0	33.3	1	Lulus
1236	0.0	0.0	0.3	0.0	0.9	0.4	0.0	36.7	57.7	85.1	83.3	0.0	16.7	1	Lulus
1240	0.0	0.0	0.7	0.0	1.7	1.1	0.0	20.1	59.7	76.7	66.7	16.7	16.7	1	Lulus
1241	0.0	0.0	0.3	0.0	0.0	1.3	0.1	36.3	53.7	87.6	27.8	55.6	16.7	1	Lulus
1245	0.0	0.0	4.6	0.0	5.5	12.6	0.8	22.4	69.0	70.1	44.4	22.2	33.3	1	Lulus
1251	0.0	0.0	0.4	0.0	1.1	0.5	0.0	59.8	69.1	80.9	66.7	0.0	33.3	1	Lulus

5. Data akademik Universitas XYZ Angkatan 2014 – 2015 bulan Desember

Instance	**A1**	**A2**	**A3**	**A4**	**A5**	**A6**	**A7**	**A8**	**A9**	**A10**	**A11**	**A12**	**A13**	**A14**	**A15**	**A16**	**A17**	**A18**	**A19**	**A20**	**A21**	**A22**	**A23**
1178	0.0	0.0	0.0	0.0	0.0	94.4	0.0	3.5	2.1	0.0	0.0	0.0	0.0	33.3	33.3	33.3	0.0	33.3	66.7	0.0	0.0	0.0	0.0
1184	0.0	0.0	0.0	0.0	0.0	98.6	0.0	0.0	1.4	0.0	0.0	0.0	0.0	0.0	0.0	100.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0
1190	0.0	0.0	0.0	0.0	0.0	11.5	0.0	0.0	88.5	0.0	18.5	31.5	13.0	22.2	0.0	14.8	52.8	0.0	35.2	12.0	0.0	0.0	0.0
1192	0.0	0.0	0.0	0.0	0.0	94.7	0.0	0.0	5.3	0.0	14.3	14.3	0.0	0.0	28.6	42.9	85.7	14.3	0.0	0.0	0.0	0.0	0.0
1196	0.0	0.0	0.0	0.0	0.0	93.7	0.0	0.0	6.4	0.0	25.0	0.0	25.0	50.0	0.0	0.0	50.0	0.0	50.0	0.0	0.0	0.0	0.0
1199	0.0	0.0	0.0	0.0	0.0	82.5	0.0	0.0	16.1	1.4	17.4	8.7	8.7	26.1	17.4	21.7	34.8	39.1	13.0	13.0	0.0	0.0	0.0
1203	0.0	0.0	0.0	0.0	0.0	98.6	0.0	0.0	1.4	0.0	0.0	0.0	0.0	0.0	50.0	50.0	50.0	0.0	0.0	50.0	0.0	0.0	0.0
1206	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1210	0.0	0.0	0.0	0.0	0.0	96.2	0.0	0.0	3.8	0.0	20.0	20.0	20.0	40.0	0.0	0.0	20.0	20.0	40.0	20.0	0.0	0.0	0.0
1211	0.0	0.0	0.0	0.0	0.0	97.9	0.0	0.0	2.1	0.0	0.0	33.3	33.3	0.0	0.0	33.3	66.7	33.3	0.0	0.0	0.0	0.0	0.0
1213	0.0	0.0	0.0	0.0	0.0	98.6	0.0	0.0	1.5	0.0	50.0	0.0	0.0	50.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0
1215	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1225	0.0	0.0	0.0	0.0	0.0	96.9	0.0	0.8	2.3	0.0	33.3	0.0	0.0	0.0	0.0	66.7	100.0	0.0	0.0	0.0	0.0	0.0	0.0
1226	0.0	0.0	0.0	0.0	0.0	97.6	0.0	0.0	2.4	0.0	0.0	0.0	0.0	33.3	33.3	33.3	33.3	0.0	33.3	33.3	0.0	0.0	0.0
1228	0.0	0.0	0.0	0.0	0.0	85.7	1.5	3.0	9.8	0.0	30.8	30.8	15.4	23.1	0.0	0.0	53.9	15.4	23.1	7.7	0.0	0.0	0.0
1229	0.0	0.0	0.0	0.0	0.0	99.2	0.0	0.0	0.8	0.0	0.0	100.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0
1230	0.0	0.0	0.0	0.0	0.0	95.2	0.0	1.6	3.2	0.0	0.0	25.0	50.0	0.0	0.0	25.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0
1232	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1233	0.0	0.0	0.0	0.0	0.0	96.4	0.0	2.2	1.5	0.0	50.0	0.0	50.0	0.0	0.0	0.0	0.0	0.0	50.0	50.0	0.0	0.0	0.0
1236	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1240	0.0	0.0	0.0	0.0	0.0	99.2	0.0	0.0	0.8	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0
1241	0.0	0.0	0.0	0.0	0.0	88.5	0.0	0.0	11.5	0.0	26.7	13.3	6.7	26.7	13.3	13.3	46.7	40.0	0.0	13.3	0.0	0.0	0.0
1245	0.0	0.0	0.0	0.0	0.0	99.3	0.0	0.0	0.8	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	100.0	0.0	0.0
1251	0.0	0.0	0.0	0.0	0.0	97.9	0.0	0.7	1.4	0.0	0.0	0.0	0.0	0.0	50.0	50.0	0.0	50.0	0.0	50.0	0.0	0.0	0.0

Instance	A24	A25	A26	A27	A28	A29	A30	A31	A32	A33	A34	A35	A36	A37	A38	A39	A40	A41	A42	Label
1178	0.0	0.0	0.1	0.0	0.3	0.1	0.0	53.4	62.9	78.3	0.0	0.0	0.0	0.0	0.0	100.0	100.0	0.0	0	Tidak Lulus
1184	0.0	0.0	0.2	0.0	0.3	0.3	0.1	51.8	62.8	78.3	0.0	0.0	0.0	0.0	0.0	100.0	100.0	0.0	0	Tidak Lulus
1190	0.0	0.0	0.0	0.0	0.0	0.0	0.0	17.4	53.8	75.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0	Tidak Lulus
1192	0.0	0.0	0.1	0.0	0.2	0.1	0.0	32.2	52.7	82.1	0.0	0.0	0.0	0.0	0.0	100.0	100.0	0.0	1	Lulus
1196	0.0	0.0	0.2	0.0	0.2	0.4	0.0	28.3	53.8	75.5	0.0	0.0	0.0	0.0	0.0	100.0	100.0	0.0	1	Lulus
1199	0.0	0.0	0.1	0.0	0.2	0.2	0.0	38.5	52.9	86.5	0.0	0.0	0.0	0.0	0.0	77.8	61.1	38.9	1	Lulus
1203	0.0	0.0	0.4	0.0	0.7	0.8	0.0	51.8	62.8	78.3	0.0	0.0	0.0	0.0	0.0	100.0	100.0	0.0	0	Tidak Lulus
1206	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0	Tidak Lulus
1210	0.0	0.0	0.2	0.0	0.2	0.5	0.0	18.8	61.7	65.9	0.0	0.0	0.0	0.0	0.0	100.0	100.0	0.0	1	Lulus
1211	0.0	0.0	0.4	0.0	0.9	0.6	0.0	53.4	62.9	78.3	0.0	0.0	0.0	0.0	0.0	100.0	100.0	0.0	0	Tidak Lulus
1213	0.0	0.0	0.1	0.0	0.1	0.1	0.0	37.7	75.9	63.1	0.0	0.0	0.0	0.0	0.0	100.0	100.0	0.0	0	Tidak Lulus
1215	0.0	0.0	0.0	0.0	0.0	0.0	0.0	33.8	58.6	70.1	0.0	0.0	0.0	0.0	0.0	100.0	100.0	0.0	0	Tidak Lulus
1225	0.0	0.0	0.0	0.0	0.0	0.0	0.2	30.5	48.9	85.5	0.0	0.0	0.0	0.0	0.0	100.0	100.0	0.0	1	Lulus
1226	0.0	0.0	0.6	0.0	1.4	0.7	0.3	28.3	53.8	75.5	0.0	0.0	0.0	0.0	0.0	100.0	83.3	16.7	1	Lulus
1228	0.0	0.0	0.0	0.0	0.1	0.0	0.0	18.8	61.7	65.9	0.0	0.0	0.0	0.0	0.0	83.3	55.6	44.4	1	Lulus
1229	0.0	0.0	0.3	0.0	0.8	0.3	0.2	17.4	53.8	75.8	0.0	0.0	0.0	0.0	0.0	100.0	100.0	0.0	1	Lulus
1230	0.0	0.0	0.9	0.0	1.2	2.1	0.3	39.6	65.5	60.3	0.0	0.0	0.0	0.0	0.0	100.0	100.0	0.0	1	Lulus
1232	0.0	0.0	0.8	0.0	1.0	2.3	0.0	51.8	62.8	78.3	0.0	0.0	0.0	0.0	0.0	100.0	100.0	0.0	1	Lulus
1233	0.0	0.0	6.7	8.0	3.6	6.5	8.9	37.7	75.9	63.1	0.0	0.0	0.0	0.0	0.0	100.0	100.0	0.0	1	Lulus
1236	0.0	0.0	0.5	0.0	1.2	0.7	0.0	32.2	52.7	82.1	0.0	0.0	0.0	0.0	0.0	100.0	100.0	0.0	1	Lulus
1240	0.0	0.0	0.8	0.0	1.7	1.4	0.0	17.4	53.8	75.8	0.0	0.0	0.0	0.0	0.0	100.0	100.0	0.0	1	Lulus
1241	0.0	0.0	0.3	0.0	0.1	1.2	0.1	30.5	48.9	85.5	0.0	0.0	0.0	0.0	0.0	100.0	83.3	16.7	1	Lulus
1245	0.0	0.0	5.2	0.0	5.7	14.4	1.5	18.8	61.7	65.9	0.0	0.0	0.0	0.0	0.0	100.0	100.0	0.0	1	Lulus
1251	0.0	0.0	0.4	0.0	1.0	0.5	0.0	51.8	62.8	78.3	0.0	0.0	0.0	0.0	0.0	100.0	100.0	0.0	1	Lulus

Lampiran 2

Atribut pada data akademik Universitas XYZ

No	Atribut Bulan Awal	Atribut Bulang Tengah Semester	Atribut Akhir Semester
1	Id_Mahasiswa	Id_Mahasiswa	Id_Mahasiswa
2	Label	Label	Label
3	Post_0_6	Post_0_6	Post_0_6
4	Post_7_12	Post_7_12	Post_7_12
5	Post_13_18	Post_13_18	Post_13_18
6	Post_18_24	Post_18_24	Post_18_24
7	Prosen_Hadir	Prosen_Hadir	Prosen_Hadir
8	Prosen_Ijin	Prosen_Ijin	Prosen_Ijin
9	Prosen_Sakit	Prosen_Sakit	Prosen_Sakit
10	Prosen_Alfa	Prosen_Alfa	Prosen_Alfa
11	Prosen_Dispensasi	Prosen_Dispensasi	Prosen_Dispensasi
12	Prosen_Alfa_Senin	Prosen_Alfa_Senin	Prosen_Alfa_Senin
13	Prosen_Alfa_Selasa	Prosen_Alfa_Selasa	Prosen_Alfa_Selasa
14	Prosen_Alfa_Rabu	Prosen_Alfa_Rabu	Prosen_Alfa_Rabu
15	Prosen_Alfa_Kamis	Prosen_Alfa_Kamis	Prosen_Alfa_Kamis
16	Prosen_Alfa_Jumat	Prosen_Alfa_Jumat	Prosen_Alfa_Jumat
17	Prosen_Alfa_Sabtu	Prosen_Alfa_Sabtu	Prosen_Alfa_Sabtu
18	Prosen_Alfa_Slot_Pagi	Prosen_Alfa_Slot_Pagi	Prosen_Alfa_Slot_Pagi
19	Prosen_Alfa_Slot_Jelang_Si ang	Prosen_Alfa_Slot_Jelang_Sia ng	Prosen_Alfa_Slot_Jelang_Sia ng
20	Prosen_Alfa_Slot_Siang	Prosen_Alfa_Slot_Siang	Prosen_Alfa_Slot_Siang
21	Prosen_Alfa_Slot_Sore	Prosen_Alfa_Slot_Sore	Prosen_Alfa_Slot_Sore
22	Prosen_Alfa_Internal	Prosen_Alfa_Internal	Prosen_Alfa_Internal
23	Frek_Pemb_Mhs	Frek_Pemb_Mhs	Frek_Pemb_Mhs
24	Frek_Akademik	Frek_Akademik	Frek_Akademik
25	Frek_Kes_Mahasiswa	Frek_Kes_Mahasiswa	Frek_Kes_Mahasiswa
26	Frek_Adm_Akademik	Frek_Adm_Akademik	Frek_Adm_Akademik
27	Durasi_Internet	Durasi_Internet	Durasi_Internet
28	Durasi_Internet_0_6	Durasi_Internet_0_6	Durasi_Internet_0_6
29	Durasi_Internet_6_12	Durasi_Internet_6_12	Durasi_Internet_6_12
30	Durasi_Internet_12_18	Durasi_Internet_12_18	Durasi_Internet_12_18
31	Durasi_Internet_18_24	Durasi_Internet_18_24	Durasi_Internet_18_24
32	Bobot_Mk_Sambung	Bobot_Mk_Sambung	Bobot_Mk_Sambung
33	Bobot_Mk_Per_Hari	Bobot_Mk_Per_Hari	Bobot_Mk_Per_Hari
34	Rerata_Jml_Hari_Kuliah_Se minggu	Rerata_Jml_Hari_Kuliah_Sem inggu	Rerata_Jml_Hari_Kuliah_Sem inggu
35		Prosen_Mk_Atas2	Jml_Ukm
36		Prosen_Mk_Bawah2	Poin_Ukm
37		Prosen_Mk_Blm_Ada	Jml_Tak
38			Poin_Tak
39			Jml_Prestasi_Mhs
40			Prosen_Sks_Lulus
41			Prosen_Jml_Mk_Atas2
42			Prosen_Jml_Mk_Bawah2

[Halaman ini sengaja dikosongkan]

Lampiran 3

Hasil Evaluasi *Classifier Discion Tree* (J48)

Hasil Evaluasi *Classifier* MWMOTE dengan Proses C*luster Average Linkage*

Agustus				
Iterasi	Precision	Recall	F-Measure	Accuracy
1	70.40%	72.00%	71.00%	72.00%
2	69.50%	71.40%	70.20%	71.41%
3	70.10%	70.70%	70.40%	70.70%
4	70.20%	71.40%	70.70%	71.41%
5	70.20%	71.00%	70.50%	70.95%
6	71.20%	74.10%	71.70%	74.09%
7	70.30%	72.10%	71.00%	72.13%
8	69.80%	70.80%	70.20%	70.75%
9	69.50%	69.90%	69.70%	69.93%
10	70.80%	73.30%	71.50%	73.30%
11	70.80%	72.30%	71.40%	72.33%
12	70.50%	72.60%	71.20%	72.56%
13	70.80%	72.50%	71.40%	72.49%
14	71.00%	73.00%	71.70%	73.05%
15	70.50%	71.90%	71.10%	71.90%

September				
Iterasi	Precision	Recall	F-Measure	Accuracy
1	73.10%	72.60%	72.80%	72.59%
2	72.00%	71.40%	71.70%	71.36%
3	74.30%	74.00%	74.10%	73.99%
4	72.70%	72.40%	72.60%	72.43%
5	73.30%	72.60%	72.90%	72.61%
6	72.70%	72.40%	72.50%	72.41%
7	72.70%	73.20%	72.90%	73.18%
8	73.70%	73.70%	73.70%	73.66%
9	73.60%	72.10%	72.70%	72.10%
10	73.50%	73.50%	73.50%	73.53%
11	73.60%	73.30%	73.40%	73.25%
12	73.60%	74.00%	73.80%	74.02%
13	72.90%	71.20%	71.90%	71.18%
14	73.20%	73.00%	73.10%	73.02%
15	73.90%	73.90%	73.90%	73.86%

Oktober				
Iterasi	Precision	Recall	F-Measure	Accuracy
1	78.00%	77.30%	77.60%	77.28%
2	78.50%	77.00%	77.60%	76.98%
3	79.60%	79.10%	79.30%	79.12%
4	79.10%	78.20%	78.60%	78.23%
5	78.30%	76.70%	77.40%	76.75%
6	78.80%	77.70%	78.10%	77.69%
7	78.00%	76.80%	77.30%	76.82%
8	78.50%	77.60%	78.00%	77.57%
9	78.40%	76.80%	77.40%	76.80%
10	79.10%	77.20%	77.90%	77.18%
11	79.70%	78.90%	79.20%	78.87%
12	78.60%	77.70%	78.10%	77.67%
13	78.70%	77.10%	77.70%	77.08%
14	78.60%	78.00%	78.30%	78.00%
15	77.60%	76.90%	77.20%	76.85%

November				
Iterasi	Precision	Recall	F-Measure	Accuracy
1	80.10%	79.10%	79.50%	79.12%
2	80.00%	78.70%	79.20%	78.71%
3	78.90%	77.30%	77.90%	77.34%
4	79.90%	78.50%	79.00%	78.46%
5	78.40%	76.90%	77.50%	76.88%
6	78.90%	77.90%	78.30%	77.90%
7	80.30%	79.30%	79.70%	79.30%
8	79.90%	78.40%	79.00%	78.43%
9	79.20%	77.70%	78.20%	77.67%
10	79.00%	77.60%	78.10%	77.57%
11	79.00%	77.70%	78.20%	77.67%
12	79.70%	78.50%	79.00%	78.54%
13	79.10%	78.10%	78.50%	78.10%
14	79.10%	78.50%	78.70%	78.46%
15	79.30%	77.90%	78.40%	77.92%

Desember				
Iterasi	Precision	Recall	F-Measure	Accuracy
1	88.30%	88.10%	88.20%	88.08%
2	88.70%	88.50%	88.60%	88.51%
3	88.90%	88.60%	88.70%	88.64%
4	88.70%	88.20%	88.40%	88.23%
5	87.60%	87.30%	87.40%	87.29%
6	87.90%	87.50%	87.70%	87.49%
7	88.10%	87.90%	88.00%	87.90%
8	87.80%	87.30%	87.50%	87.32%
9	87.80%	87.30%	87.50%	87.29%
10	88.20%	87.90%	88.00%	87.93%
11	88.50%	88.10%	88.30%	88.13%
12	88.60%	88.40%	88.50%	88.44%
13	87.30%	86.90%	87.10%	86.88%
14	88.90%	88.60%	88.70%	88.64%
15	88.60%	88.50%	88.60%	88.49%

Hasil Evaluasi *Classifier* MWMOTE dengan Proses C*luster Complete Linkage*

Agustus				
Iterasi	Precision	Recall	F-Measure	Accuracy
1	70.90%	72.00%	71.40%	72.03%
2	70.30%	70.80%	70.50%	70.78%
3	70.40%	72.10%	71.00%	72.10%
4	70.20%	71.30%	70.70%	71.29%
5	69.80%	70.40%	70.10%	70.44%
6	70.20%	69.90%	70.00%	69.93%
7	70.20%	70.80%	70.50%	70.80%
8	70.40%	72.40%	71.10%	72.38%
9	70.40%	72.20%	71.10%	72.21%
10	70.30%	70.20%	70.20%	70.16%
11	69.80%	72.00%	70.60%	72.03%
12	70.60%	71.40%	71.00%	71.44%
13	69.90%	71.40%	70.50%	71.39%
14	70.20%	70.60%	70.40%	70.65%
15	70.40%	71.60%	70.90%	71.62%

September				
Iterasi	Precision	Recall	F-Measure	Accuracy
1	73.10%	72.70%	72.90%	72.69%
2	73.90%	73.70%	73.80%	73.69%
3	72.80%	72.30%	72.50%	72.33%
4	74.00%	74.10%	74.10%	74.15%
5	73.60%	72.50%	73.00%	72.49%
6	73.90%	73.70%	73.80%	73.69%
7	73.30%	72.80%	73.00%	72.82%
8	73.00%	72.60%	72.80%	72.56%
9	73.60%	73.60%	73.60%	73.61%
10	73.80%	73.40%	73.60%	73.43%
11	73.40%	73.60%	73.50%	73.61%
12	73.90%	74.30%	74.10%	74.30%
13	73.30%	72.90%	73.10%	72.87%
14	72.20%	72.10%	72.20%	72.13%
15	72.80%	72.70%	72.70%	72.66%

Oktober				
Iterasi	Precision	Recall	F-Measure	Accuracy
1	79.00%	77.70%	78.20%	77.72%
2	78.80%	78.30%	78.50%	78.28%
3	78.50%	77.60%	78.00%	77.59%
4	78.90%	78.70%	78.80%	78.74%
5	78.30%	77.30%	77.70%	77.26%
6	78.30%	77.60%	77.90%	77.59%
7	79.60%	78.60%	79.00%	78.59%
8	78.00%	76.60%	77.20%	76.62%
9	78.20%	77.00%	77.50%	76.95%
10	77.80%	76.50%	77.00%	76.52%
11	78.30%	78.00%	78.20%	78.02%
12	78.90%	77.50%	78.00%	77.46%
13	79.00%	77.90%	78.30%	77.87%
14	79.40%	78.00%	78.50%	77.97%
15	79.20%	77.70%	78.20%	77.67%

November				
Iterasi	Precision	Recall	F-Measure	Accuracy
1	79.20%	77.70%	78.30%	77.72%
2	80.20%	78.70%	79.20%	78.66%
3	79.70%	78.40%	78.90%	78.36%
4	80.30%	78.70%	79.30%	78.74%
5	79.60%	78.50%	78.90%	78.48%
6	79.00%	77.50%	78.10%	77.54%
7	79.80%	78.50%	79.00%	78.48%
8	79.20%	78.50%	78.80%	78.51%
9	79.30%	78.00%	78.50%	77.97%
10	79.20%	77.60%	78.20%	77.62%
11	79.00%	78.40%	78.70%	78.38%
12	80.00%	79.00%	79.40%	78.97%
13	79.40%	77.80%	78.40%	77.85%
14	79.00%	77.70%	78.20%	77.74%
15	80.00%	77.60%	78.40%	77.59%

Desember				
Iterasi	Precision	Recall	F-Measure	Accuracy
1	88.40%	87.70%	87.90%	87.67%
2	87.80%	87.50%	87.60%	87.47%
3	88.10%	87.60%	87.80%	87.65%
4	88.70%	88.60%	88.60%	88.59%
5	88.40%	88.10%	88.20%	88.13%
6	88.20%	87.90%	88.00%	87.88%
7	88.90%	88.70%	88.80%	88.72%
8	88.30%	87.90%	88.00%	87.88%
9	87.80%	87.30%	87.50%	87.32%
10	88.50%	88.30%	88.40%	88.28%
11	88.20%	87.80%	87.90%	87.77%
12	87.50%	87.10%	87.20%	87.06%
13	88.00%	87.60%	87.70%	87.57%
14	88.50%	88.30%	88.40%	88.31%
15	88.30%	87.90%	88.10%	87.90%

Hasil Evaluasi *Classifier* MWMOTE dengan Proses C*luster Single Linkage*

Agustus				
Iterasi	Precision	Recall	F-Measure	Accuracy
1	69.80%	70.20%	70.00%	70.24%
2	69.50%	71.20%	70.20%	71.18%
3	70.50%	71.60%	70.90%	71.57%
4	70.80%	72.80%	71.50%	72.79%
5	70.70%	72.40%	71.30%	72.43%
6	71.20%	73.40%	71.90%	73.35%
7	69.80%	71.40%	70.50%	71.36%
8	71.20%	72.70%	71.80%	72.72%
9	70.30%	71.40%	70.80%	71.44%
10	70.60%	71.60%	71.10%	71.62%
11	69.90%	70.70%	70.30%	70.72%
12	69.50%	70.80%	70.00%	70.75%
13	69.30%	70.40%	69.80%	70.37%
14	69.80%	70.30%	70.10%	70.34%
15	69.60%	71.50%	70.30%	71.54%

September				
Iterasi	Precision	Recall	F-Measure	Accuracy
1	73.00%	73.10%	73.10%	73.07%
2	73.30%	72.90%	73.10%	72.87%
3	74.10%	74.40%	74.30%	74.37%
4	72.40%	72.30%	72.30%	72.26%
5	73.00%	72.80%	72.90%	72.77%
6	73.30%	73.00%	73.10%	73.00%
7	74.00%	74.00%	74.00%	73.99%
8	73.50%	73.90%	73.70%	73.92%
9	73.80%	73.70%	73.70%	73.69%
10	73.00%	72.70%	72.90%	72.74%
11	74.10%	73.70%	73.90%	73.71%
12	73.50%	73.60%	73.50%	73.58%
13	72.40%	72.80%	72.60%	72.79%
14	73.80%	73.80%	73.80%	73.76%
15	73.70%	74.00%	73.80%	73.97%

Oktober				
Iterasi	Precision	Recall	F-Measure	Accuracy
1	79.00%	78.30%	78.60%	78.31%
2	78.90%	77.80%	78.20%	77.79%
3	78.10%	76.50%	77.10%	76.54%
4	78.10%	77.50%	77.70%	77.46%
5	79.20%	78.20%	78.60%	78.23%
6	78.80%	78.10%	78.40%	78.05%
7	78.70%	77.90%	78.20%	77.87%
8	78.70%	77.40%	77.90%	77.36%
9	78.80%	76.70%	77.50%	76.75%
10	79.10%	77.80%	78.30%	77.82%
11	79.70%	79.30%	79.50%	79.33%
12	78.20%	77.10%	77.50%	77.08%
13	79.30%	79.00%	79.10%	78.97%
14	78.90%	77.80%	78.20%	77.77%
15	79.60%	78.30%	78.80%	78.31%

November				
Iterasi	Precision	Recall	F-Measure	Accuracy
1	78.90%	77.30%	77.90%	77.34%
2	78.60%	77.80%	78.20%	77.82%
3	80.70%	79.80%	80.10%	79.79%
4	79.40%	77.90%	78.40%	77.87%
5	78.30%	76.70%	77.30%	76.70%
6	79.90%	78.40%	78.90%	78.36%
7	79.50%	78.50%	78.90%	78.46%
8	79.60%	78.30%	78.80%	78.33%
9	79.10%	77.80%	78.30%	77.85%
10	78.70%	77.20%	77.80%	77.21%
11	79.80%	78.30%	78.90%	78.33%
12	79.90%	78.50%	79.00%	78.48%
13	79.60%	78.70%	79.00%	78.66%
14	80.50%	79.70%	80.00%	79.71%
15	79.70%	78.40%	78.90%	78.38%

Desember				
Iterasi	Precision	Recall	F-Measure	Accuracy
1	88.60%	88.60%	88.60%	88.57%
2	88.40%	88.00%	88.20%	88.03%
3	88.20%	87.90%	88.00%	87.88%
4	87.70%	87.20%	87.40%	87.19%
5	87.60%	87.00%	87.20%	87.03%
6	88.30%	87.80%	88.00%	87.80%
7	88.30%	88.00%	88.10%	88.00%
8	88.40%	88.10%	88.20%	88.08%
9	88.10%	87.50%	87.70%	87.52%
10	87.50%	87.10%	87.30%	87.11%
11	88.40%	88.00%	88.10%	87.98%
12	88.20%	87.70%	87.90%	87.75%
13	88.50%	88.20%	88.30%	88.21%
14	88.70%	88.30%	88.40%	88.28%
15	88.00%	87.80%	87.90%	87.77%

Lampiran 4

Hasil Uji Statistik *One Way* ANOVA data akademik Universitas XYZ angkatan 2014 dan 2015

HASIL UJI STATISTIK BULAN DESEMBER

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	*Statistic*	df	Sig.
HASIL UJI BULAN DESEMBER	.078	45	.200*	.973	45	.383

**. This is a lower bound of the true significance. a. Lilliefors Significance Correction*

Test of Homogeneity of Variances

Levene Statistic	df1	df2	Sig.
1.082	2	42	.348

Test of ANOVA

	Sum of Squares	df	*Mean Square*	F	Sig.
Between Groups	.000	2	.000	.291	.749
Within Groups	.001	42	.000		
Total	.001	44			

***Test of* Duncan[a]**

METODE *CLUSTERING*	N	*Subset for alpha* = 0.05 1
SINGLE	15	.87813520
COMPLETE	15	.87879880
AVERAGE	15	.87951340
Sig.		.478

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 15.000.

HASIL UJI STATISTIK BULAN NOVEMBER

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	*Statistic*	df	Sig.
HASIL UJI BULAN NOVEMBER	.126	45	.070	.974	45	.401

a. Lilliefors Significance Correction

Test of Homogeneity of Variances

Levene Statistic	df1	df2	Sig.
.811	2	42	.451

***Test of* ANOVA**

	Sum of Squares	df	*Mean Square*	F	Sig.
Between Groups	.000	2	.000	.055	.946
Within Groups	.002	42	.000		
Total	.002	44			

***Test of* Duncan[a]**

METODE CLUSTERING	N	*Subset for alpha* = 0.05 1
AVERAGE	15	.78136820
COMPLETE	15	.78174233
SINGLE	15	.78218480
Sig.		.757

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 15.000.

HASIL UJI STATISTIK BULAN OKTOBER

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	*Statistic*	df	Sig.
HASIL UJI BULAN OKTOBER	.065	45	.200*	.971	45	.327

**. This is a lower bound of the true significance.*

a. Lilliefors Significance Correction

Test of Homogeneity of Variances

Levene Statistic	df1	df2	Sig.
.257	2	42	.774

***Test of* ANOVA**

	Sum of Squares	df	*Mean Square*	F	Sig.
Between Groups	.000	2	.000	.738	.484
Within Groups	.002	42	.000		
Total	.002	44			

***Test of* Duncan[a]**

METODE CLUSTERING	N	*Subset for alpha* = 0.05 1
AVERAGE	15	.77525947
COMPLETE	15	.77656967
SINGLE	15	.77842447
Sig.		.261

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 15.000.

HASIL UJI STATISTIK BULAN SEPTEMBER

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	*Statistic*	df	Sig.
HASIL UJI BULAN SEPTEMBER	.129	45	.058	.959	45	.109

a. Lilliefors Significance Correction

Test of Homogeneity of Variances

Levene Statistic	df1	df2	Sig.
1.129	2	42	.333

***Test of* ANOVA**

	Sum of Squares	df	*Mean Square*	F	Sig.
Between Groups	.000	2	.000	1.617	.211
Within Groups	.002	42	.000		
Total	.002	44			

***Test of* Duncan[a]**

METODE CLUSTERING	N	*Subset for alpha* = 0.05 1
AVERAGE	15	.72880720
COMPLETE	15	.73134253
SINGLE	15	.73365667
Sig.		.096

Means for groups in homogeneous subsets are displayed.
a. Uses Harmonic Mean Sample Size = 15.000.

HASIL UJI STATISTIK BULAN AGUSTUS

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	*Statistic*	df	Sig.
HASIL UJI BULAN AGUSTUS	.096	45	.200*	.977	45	.500

**. This is a lower bound of the true significance. a. Lilliefors Significance Correction*

Test of Homogeneity of Variances

Levene Statistic	df1	df2	Sig.
.528	2	42	.594

***Test of* ANOVA**

	Sum of Squares	df	*Mean Square*	F	Sig.
Between Groups	.000	2	.000	1.796	.178
Within Groups	.004	42	.000		
Total	.004	44			

***Test of* Duncan[a]**

METODE CLUSTERING	N	*Subset for alpha* = 0.05 1
COMPLETE	15	.71282960
SINGLE	15	.71495660
AVERAGE	15	.71934653
Sig.		.085

Means for groups in homogeneous subsets are displayed. a. Uses Harmonic Mean Sample Size = 15.000.

[Halaman ini sengaja dikosongkan]

Lampiran 5

Hasil klassifikasi *descion tree* (J48) dataset 10 UCI

	AVERAGE LINKAGE				COMPLETE LINKAGE				SINGLE LINKAGE			
ABALONE	Precision	Recall	F-Measure	Accuracy	Precision	Recall	F-Measure	Accuracy	Precision	Recall	F-Measure	Accuracy
1	98.60%	98.30%	98.40%	98.27%	98.80%	98.40%	98.50%	98.45%	96.60%	93.40%	94.50%	93.44%
2	97.90%	96.50%	96.90%	96.55%	98.50%	97.90%	98.10%	97.93%	96.10%	90.30%	92.30%	90.33%
3	98.70%	98.40%	98.50%	98.45%	98.40%	97.80%	97.90%	97.75%	97.60%	96.20%	96.60%	96.20%
4	98.70%	98.30%	98.40%	98.27%	98.20%	97.60%	97.80%	97.58%	97.60%	96.20%	96.60%	96.20%
5	98.20%	97.40%	97.60%	97.41%	98.80%	98.40%	98.50%	98.45%	97.20%	96.70%	96.90%	96.72%
6	98.60%	98.30%	98.40%	98.27%	98.80%	98.40%	98.50%	98.45%	96.60%	93.40%	94.50%	93.44%
7	97.90%	96.50%	96.90%	96.55%	98.50%	97.90%	98.10%	97.93%	96.10%	90.30%	92.30%	90.33%
8	98.70%	98.40%	98.50%	98.45%	98.40%	97.80%	97.90%	97.75%	97.60%	96.20%	96.60%	96.20%
9	98.70%	98.30%	98.40%	98.27%	98.20%	97.60%	97.80%	97.58%	97.60%	96.20%	96.60%	96.20%
10	98.20%	97.40%	97.60%	97.41%	98.80%	98.40%	98.50%	98.45%	97.20%	96.70%	96.90%	96.72%
11	98.60%	98.30%	98.40%	98.27%	98.80%	98.40%	98.50%	98.45%	96.60%	93.40%	94.50%	93.44%
12	97.90%	96.50%	96.90%	96.55%	98.50%	97.90%	98.10%	97.93%	96.10%	90.30%	92.30%	90.33%
13	98.70%	98.40%	98.50%	98.45%	98.40%	97.80%	97.90%	97.75%	97.60%	96.20%	96.60%	96.20%
14	98.70%	98.30%	98.40%	98.27%	98.20%	97.60%	97.80%	97.58%	97.60%	96.20%	96.60%	96.20%
15	98.20%	97.40%	97.60%	97.41%	98.80%	98.40%	98.50%	98.45%	97.20%	96.70%	96.90%	96.72%
BREAS	Precision	Recall	F-Measure	Accuracy	Precision	Recall	F-Measure	Accuracy	Precision	Recall	F-Measure	Accuracy
1	88.70%	85.20%	85.60%	85.19%	91.70%	88.90%	89.20%	88.89%	98.80%	98.80%	98.80%	98.77%
2	90.50%	86.40%	86.90%	86.42%	91.10%	87.70%	88.00%	87.65%	98.80%	98.80%	98.80%	98.77%
3	97.50%	97.50%	97.50%	97.53%	100.00%	100.00%	100.00%	100.00%	88.80%	82.70%	83.30%	82.72%
4	90.50%	86.40%	86.90%	86.42%	95.70%	95.10%	95.10%	95.06%	88.80%	82.70%	83.30%	82.72%
5	91.10%	87.70%	88.00%	87.65%	95.70%	95.10%	95.10%	95.06%	88.30%	81.50%	82.10%	81.48%

6	88.70%	85.20%	85.60%	85.19%		91.70%	88.90%	89.20%	88.89%		98.80%	98.80%	98.80%	98.77%
7	90.50%	86.40%	86.90%	86.42%		91.10%	87.70%	88.00%	87.65%		98.80%	98.80%	98.80%	98.77%
8	97.50%	97.50%	97.50%	97.53%		100.00%	100.00%	100.00%	100.00%		88.80%	82.70%	83.30%	82.72%
9	90.50%	86.40%	86.90%	86.42%		95.70%	95.10%	95.10%	95.06%		88.80%	82.70%	83.30%	82.72%
10	91.10%	87.70%	88.00%	87.65%		95.70%	95.10%	95.10%	95.06%		88.30%	81.50%	82.10%	81.48%
11	88.70%	85.20%	85.60%	85.19%		91.70%	88.90%	89.20%	88.89%		98.80%	98.80%	98.80%	98.77%
12	90.50%	86.40%	86.90%	86.42%		91.10%	87.70%	88.00%	87.65%		98.80%	98.80%	98.80%	98.77%
13	97.50%	97.50%	97.50%	97.53%		100.00%	100.00%	100.00%	100.00%		88.80%	82.70%	83.30%	82.72%
14	90.50%	86.40%	86.90%	86.42%		95.70%	95.10%	95.10%	95.06%		88.80%	82.70%	83.30%	82.72%
15	91.10%	87.70%	88.00%	87.65%		95.70%	95.10%	95.10%	95.06%		88.30%	81.50%	82.10%	81.48%
ECOLI	Precision	Recall	F-Measure	Accuracy		Precision	Recall	F-Measure	Accuracy		Precision	Recall	F-Measure	Accuracy
1	94.70%	93.00%	93.30%	92.97%		94.90%	93.40%	93.70%	93.36%		95.40%	94.10%	94.40%	94.14%
2	94.70%	93.00%	93.30%	92.97%		94.20%	92.60%	92.90%	92.58%		94.70%	93.40%	93.60%	93.36%
3	96.40%	95.70%	95.80%	95.70%		94.30%	92.20%	92.60%	92.19%		95.40%	94.10%	94.40%	94.14%
4	94.70%	93.00%	93.30%	92.97%		94.70%	93.00%	93.30%	92.97%		94.70%	93.80%	94.00%	93.75%
5	95.40%	94.10%	94.40%	94.14%		94.70%	93.00%	93.30%	92.97%		97.10%	96.90%	96.90%	96.88%
6	94.70%	93.00%	93.30%	92.97%		94.90%	93.40%	93.70%	93.36%		95.40%	94.10%	94.40%	94.14%
7	94.70%	93.00%	93.30%	92.97%		94.20%	92.60%	92.90%	92.58%		94.70%	93.40%	93.60%	93.36%
8	96.40%	95.70%	95.80%	95.70%		94.30%	92.20%	92.60%	92.19%		95.40%	94.10%	94.40%	94.14%
9	94.70%	93.00%	93.30%	92.97%		94.70%	93.00%	93.30%	92.97%		94.70%	93.80%	94.00%	93.75%
10	95.40%	94.10%	94.40%	94.14%		94.70%	93.00%	93.30%	92.97%		97.10%	96.90%	96.90%	96.88%
11	94.70%	93.00%	93.30%	92.97%		94.90%	93.40%	93.70%	93.36%		95.40%	94.10%	94.40%	94.14%
12	94.70%	93.00%	93.30%	92.97%		94.20%	92.60%	92.90%	92.58%		94.70%	93.40%	93.60%	93.36%
13	96.40%	95.70%	95.80%	95.70%		94.30%	92.20%	92.60%	92.19%		95.40%	94.10%	94.40%	94.14%
14	94.70%	93.00%	93.30%	92.97%		94.70%	93.00%	93.30%	92.97%		94.70%	93.80%	94.00%	93.75%
15	95.40%	94.10%	94.40%	94.14%		94.70%	93.00%	93.30%	92.97%		97.10%	96.90%	96.90%	96.88%

GLASS	Precision	Recall	F-Measure	Accuracy		Precision	Recall	F-Measure	Accuracy		Precision	Recall	F-Measure	Accuracy
1	98.10%	98.10%	98.10%	98.11%		98.70%	98.70%	98.70%	98.74%		98.10%	98.10%	98.10%	98.11%
2	98.70%	98.70%	98.70%	98.74%		99.10%	99.10%	99.10%	99.08%		98.80%	98.70%	98.80%	98.74%
3	98.10%	98.10%	98.10%	98.11%		100.00%	100.00%	100.00%	100.00%		98.10%	98.10%	98.10%	98.11%
4	99.40%	99.40%	99.40%	99.37%		98.10%	98.10%	98.10%	98.11%		99.40%	99.40%	99.40%	99.37%
5	97.50%	97.50%	97.50%	97.48%		98.80%	98.70%	98.80%	98.74%		99.40%	99.40%	99.40%	99.37%
6	98.10%	98.10%	98.10%	98.11%		98.70%	98.70%	98.70%	98.74%		98.10%	98.10%	98.10%	98.11%
7	98.70%	98.70%	98.70%	98.74%		99.10%	99.10%	99.10%	99.08%		98.80%	98.70%	98.80%	98.74%
8	98.10%	98.10%	98.10%	98.11%		100.00%	100.00%	100.00%	100.00%		98.10%	98.10%	98.10%	98.11%
9	99.40%	99.40%	99.40%	99.37%		98.10%	98.10%	98.10%	98.11%		99.40%	99.40%	99.40%	99.37%
10	97.50%	97.50%	97.50%	97.48%		98.80%	98.70%	98.80%	98.74%		99.40%	99.40%	99.40%	99.37%
11	98.10%	98.10%	98.10%	98.11%		98.70%	98.70%	98.70%	98.74%		98.10%	98.10%	98.10%	98.11%
12	98.70%	98.70%	98.70%	98.74%		99.10%	99.10%	99.10%	99.08%		98.80%	98.70%	98.80%	98.74%
13	98.10%	98.10%	98.10%	98.11%		100.00%	100.00%	100.00%	100.00%		98.10%	98.10%	98.10%	98.11%
14	99.40%	99.40%	99.40%	99.37%		98.10%	98.10%	98.10%	98.11%		99.40%	99.40%	99.40%	99.37%
15	97.50%	97.50%	97.50%	97.48%		98.80%	98.70%	98.80%	98.74%		99.40%	99.40%	99.40%	99.37%
LIBRA	Precision	Recall	F-Measure	Accuracy		Precision	Recall	F-Measure	Accuracy		Precision	Recall	F-Measure	Accuracy
1	99.30%	99.30%	99.30%	99.27%		99.00%	98.90%	98.90%	98.91%		98.50%	98.50%	98.50%	98.54%
2	99.30%	99.30%	99.30%	99.27%		98.60%	98.50%	98.60%	98.54%		97.90%	97.80%	97.80%	97.81%
3	99.60%	99.60%	99.60%	99.64%		99.60%	99.60%	99.60%	99.64%		99.30%	99.30%	99.30%	99.27%
4	98.90%	98.90%	98.90%	98.91%		98.90%	98.90%	98.90%	98.91%		99.30%	99.30%	99.30%	99.27%
5	98.20%	98.20%	98.20%	98.18%		99.30%	99.30%	99.30%	99.27%		98.90%	98.90%	98.90%	98.91%
6	99.30%	99.30%	99.30%	99.27%		99.00%	98.90%	98.90%	98.91%		98.50%	98.50%	98.50%	98.54%
7	99.30%	99.30%	99.30%	99.27%		98.60%	98.50%	98.60%	98.54%		97.90%	97.80%	97.80%	97.81%
8	99.60%	99.60%	99.60%	99.64%		99.60%	99.60%	99.60%	99.64%		99.30%	99.30%	99.30%	99.27%
9	98.90%	98.90%	98.90%	98.91%		98.90%	98.90%	98.90%	98.91%		99.30%	99.30%	99.30%	99.27%

10	98.20%	98.20%	98.20%	98.18%		99.30%	99.30%	99.30%	99.27%		98.90%	98.90%	98.90%	98.91%
11	99.30%	99.30%	99.30%	99.27%		99.00%	98.90%	98.90%	98.91%		98.50%	98.50%	98.50%	98.54%
12	99.30%	99.30%	99.30%	99.27%		98.60%	98.50%	98.60%	98.54%		97.90%	97.80%	97.80%	97.81%
13	99.60%	99.60%	99.60%	99.64%		99.60%	99.60%	99.60%	99.64%		99.30%	99.30%	99.30%	99.27%
14	98.90%	98.90%	98.90%	98.91%		98.90%	98.90%	98.90%	98.91%		99.30%	99.30%	99.30%	99.27%
15	98.20%	98.20%	98.20%	98.18%		99.30%	99.30%	99.30%	99.27%		98.90%	98.90%	98.90%	98.91%
OCR	Precision	Recall	F-Measure	Accuracy		Precision	Recall	F-Measure	Accuracy		Precision	Recall	F-Measure	Accuracy
1	100.00%	100.00%	100.00%	99.97%		99.90%	99.90%	99.90%	99.90%		99.90%	99.90%	99.90%	99.90%
2	99.90%	99.90%	99.90%	99.90%		99.90%	99.90%	99.90%	99.93%		100.00%	100.00%	100.00%	100.00%
3	99.90%	99.90%	99.90%	99.90%		100.00%	100.00%	100.00%	99.97%		99.80%	99.80%	99.80%	99.84%
4	99.80%	99.80%	99.80%	99.84%		100.00%	100.00%	100.00%	100.00%		99.90%	99.90%	99.90%	99.93%
5	100.00%	100.00%	100.00%	100.00%		100.00%	100.00%	100.00%	99.97%		99.90%	99.90%	99.90%	99.90%
6	100.00%	100.00%	100.00%	99.97%		99.90%	99.90%	99.90%	99.90%		99.90%	99.90%	99.90%	99.90%
7	99.90%	99.90%	99.90%	99.90%		99.90%	99.90%	99.90%	99.93%		100.00%	100.00%	100.00%	100.00%
8	99.90%	99.90%	99.90%	99.90%		100.00%	100.00%	100.00%	99.97%		99.80%	99.80%	99.80%	99.84%
9	99.80%	99.80%	99.80%	99.84%		100.00%	100.00%	100.00%	100.00%		99.90%	99.90%	99.90%	99.93%
10	100.00%	100.00%	100.00%	100.00%		100.00%	100.00%	100.00%	99.97%		99.90%	99.90%	99.90%	99.90%
11	100.00%	100.00%	100.00%	99.97%		99.90%	99.90%	99.90%	99.90%		99.90%	99.90%	99.90%	99.90%
12	99.90%	99.90%	99.90%	99.90%		99.90%	99.90%	99.90%	99.93%		100.00%	100.00%	100.00%	100.00%
13	99.90%	99.90%	99.90%	99.90%		100.00%	100.00%	100.00%	99.97%		99.80%	99.80%	99.80%	99.84%
14	99.80%	99.80%	99.80%	99.84%		100.00%	100.00%	100.00%	100.00%		99.90%	99.90%	99.90%	99.93%
15	100.00%	100.00%	100.00%	100.00%		100.00%	100.00%	100.00%	99.97%		99.90%	99.90%	99.90%	99.90%
ROBOT	Precision	Recall	F-Measure	Accuracy		Precision	Recall	F-Measure	Accuracy		Precision	Recall	F-Measure	Accuracy
1	99.90%	99.90%	99.90%	99.89%		99.90%	99.90%	99.90%	99.86%		99.80%	99.80%	99.80%	99.79%
2	99.80%	99.80%	99.80%	99.84%		99.80%	99.80%	99.80%	99.84%		99.80%	99.80%	99.80%	99.84%
3	99.90%	99.90%	99.90%	99.91%		99.90%	99.90%	99.90%	99.86%		99.80%	99.80%	99.80%	99.82%

4	99.90%	99.90%	99.90%	99.86%		99.90%	99.90%	99.90%	99.91%		99.90%	99.90%	99.90%	99.86%
5	99.90%	99.90%	99.90%	99.89%		99.90%	99.90%	99.90%	99.93%		99.80%	99.80%	99.80%	99.77%
6	99.90%	99.90%	99.90%	99.89%		99.90%	99.90%	99.90%	99.86%		99.80%	99.80%	99.80%	99.79%
7	99.80%	99.80%	99.80%	99.84%		99.80%	99.80%	99.80%	99.84%		99.80%	99.80%	99.80%	99.84%
8	99.90%	99.90%	99.90%	99.91%		99.90%	99.90%	99.90%	99.86%		99.80%	99.80%	99.80%	99.82%
9	99.90%	99.90%	99.90%	99.86%		99.90%	99.90%	99.90%	99.91%		99.90%	99.90%	99.90%	99.86%
10	99.90%	99.90%	99.90%	99.89%		99.90%	99.90%	99.90%	99.93%		99.80%	99.80%	99.80%	99.77%
11	99.90%	99.90%	99.90%	99.89%		99.90%	99.90%	99.90%	99.86%		99.80%	99.80%	99.80%	99.79%
12	99.80%	99.80%	99.80%	99.84%		99.80%	99.80%	99.80%	99.84%		99.80%	99.80%	99.80%	99.84%
13	99.90%	99.90%	99.90%	99.91%		99.90%	99.90%	99.90%	99.86%		99.80%	99.80%	99.80%	99.82%
14	99.90%	99.90%	99.90%	99.86%		99.90%	99.90%	99.90%	99.91%		99.90%	99.90%	99.90%	99.86%
15	99.90%	99.90%	99.90%	99.89%		99.90%	99.90%	99.90%	99.93%		99.80%	99.80%	99.80%	99.77%
STATIGAMER	Precision	Recall	F-Measure	Accuracy		Precision	Recall	F-Measure	Accuracy		Precision	Recall	F-Measure	Accuracy
1	99.90%	99.90%	99.90%	99.89%		100.00%	100.00%	100.00%	100.00%		99.80%	99.80%	99.80%	99.84%
2	99.90%	99.90%	99.90%	99.89%		100.00%	100.00%	100.00%	100.00%		99.90%	99.90%	99.90%	99.95%
3	99.90%	99.90%	99.90%	99.89%		100.00%	100.00%	100.00%	100.00%		99.80%	99.80%	99.80%	99.84%
4	99.90%	99.90%	99.90%	99.89%		100.00%	100.00%	100.00%	100.00%		99.90%	99.90%	99.90%	99.89%
5	99.90%	99.90%	99.90%	99.89%		100.00%	100.00%	100.00%	100.00%		99.90%	99.90%	99.90%	99.95%
6	99.90%	99.90%	99.90%	99.89%		100.00%	100.00%	100.00%	100.00%		99.80%	99.80%	99.80%	99.84%
7	99.90%	99.90%	99.90%	99.89%		100.00%	100.00%	100.00%	100.00%		99.90%	99.90%	99.90%	99.95%
8	99.90%	99.90%	99.90%	99.89%		100.00%	100.00%	100.00%	100.00%		99.80%	99.80%	99.80%	99.84%
9	99.90%	99.90%	99.90%	99.89%		100.00%	100.00%	100.00%	100.00%		99.90%	99.90%	99.90%	99.89%
10	99.90%	99.90%	99.90%	99.89%		100.00%	100.00%	100.00%	100.00%		99.90%	99.90%	99.90%	99.95%
11	99.90%	99.90%	99.90%	99.89%		100.00%	100.00%	100.00%	100.00%		99.80%	99.80%	99.80%	99.84%
12	99.90%	99.90%	99.90%	99.89%		100.00%	100.00%	100.00%	100.00%		99.90%	99.90%	99.90%	99.95%
13	99.90%	99.90%	99.90%	99.89%		100.00%	100.00%	100.00%	100.00%		99.80%	99.80%	99.80%	99.84%

14	99.90%	99.90%	99.90%	99.89%		100.00%	100.00%	100.00%	100.00%		99.90%	99.90%	99.90%	99.89%
15	99.90%	99.90%	99.90%	99.89%		100.00%	100.00%	100.00%	100.00%		99.90%	99.90%	99.90%	99.95%
WIN	Precision	Recall	F-Measure	Accuracy		Precision	Recall	F-Measure	Accuracy		Precision	Recall	F-Measure	Accuracy
1	96.70%	96.30%	96.30%	96.27%		99.30%	99.30%	99.30%	99.25%		99.30%	99.30%	99.30%	99.25%
2	96.70%	96.30%	96.30%	96.27%		99.30%	99.30%	99.30%	99.25%		99.30%	99.30%	99.30%	99.25%
3	99.30%	99.30%	99.30%	99.25%		100.00%	100.00%	100.00%	100.00%		99.30%	99.30%	99.30%	99.25%
4	99.30%	99.30%	99.30%	99.25%		98.50%	98.50%	98.50%	98.51%		96.70%	96.30%	96.30%	96.27%
5	99.30%	99.30%	99.30%	99.25%		99.30%	99.30%	99.30%	99.25%		99.30%	99.30%	99.30%	99.25%
6	96.70%	96.30%	96.30%	96.27%		99.30%	99.30%	99.30%	99.25%		99.30%	99.30%	99.30%	99.25%
7	96.70%	96.30%	96.30%	96.27%		99.30%	99.30%	99.30%	99.25%		99.30%	99.30%	99.30%	99.25%
8	99.30%	99.30%	99.30%	99.25%		100.00%	100.00%	100.00%	100.00%		99.30%	99.30%	99.30%	99.25%
9	99.30%	99.30%	99.30%	99.25%		98.50%	98.50%	98.50%	98.51%		96.70%	96.30%	96.30%	96.27%
10	99.30%	99.30%	99.30%	99.25%		99.30%	99.30%	99.30%	99.25%		99.30%	99.30%	99.30%	99.25%
11	96.70%	96.30%	96.30%	96.27%		99.30%	99.30%	99.30%	99.25%		99.30%	99.30%	99.30%	99.25%
12	96.70%	96.30%	96.30%	96.27%		99.30%	99.30%	99.30%	99.25%		99.30%	99.30%	99.30%	99.25%
13	99.30%	99.30%	99.30%	99.25%		100.00%	100.00%	100.00%	100.00%		99.30%	99.30%	99.30%	99.25%
14	99.30%	99.30%	99.30%	99.25%		98.50%	98.50%	98.50%	98.51%		96.70%	96.30%	96.30%	96.27%
15	99.30%	99.30%	99.30%	99.25%		99.30%	99.30%	99.30%	99.25%		99.30%	99.30%	99.30%	99.25%
YEAS	Precision	Recall	F-Measure	Accuracy		Precision	Recall	F-Measure	Accuracy		Precision	Recall	F-Measure	Accuracy
1	92.50%	91.80%	92.00%	91.78%		94.10%	93.50%	93.70%	93.47%		92.90%	91.60%	92.00%	91.61%
2	93.70%	93.20%	93.40%	93.22%		94.60%	93.70%	94.00%	93.73%		93.30%	93.10%	93.20%	93.05%
3	94.10%	94.10%	94.10%	94.07%		94.80%	94.30%	94.50%	94.32%		90.90%	88.00%	88.70%	87.97%
4	93.10%	91.90%	92.30%	91.95%		93.90%	93.30%	93.50%	93.31%		92.90%	91.90%	92.20%	91.95%
5	94.20%	94.00%	94.10%	93.98%		94.40%	93.80%	94.00%	93.81%		92.70%	91.60%	91.90%	91.61%
6	92.50%	91.80%	92.00%	91.78%		94.10%	93.50%	93.70%	93.47%		92.90%	91.60%	92.00%	91.61%
7	93.70%	93.20%	93.40%	93.22%		94.60%	93.70%	94.00%	93.73%		93.30%	93.10%	93.20%	93.05%

8	94.10%	94.10%	94.10%	94.07%		94.80%	94.30%	94.50%	94.32%		90.90%	88.00%	88.70%	87.97%
9	93.10%	91.90%	92.30%	91.95%		93.90%	93.30%	93.50%	93.31%		92.90%	91.90%	92.20%	91.95%
10	94.20%	94.00%	94.10%	93.98%		94.40%	93.80%	94.00%	93.81%		92.70%	91.60%	91.90%	91.61%
11	92.50%	91.80%	92.00%	91.78%		94.10%	93.50%	93.70%	93.47%		92.90%	91.60%	92.00%	91.61%
12	93.70%	93.20%	93.40%	93.22%		94.60%	93.70%	94.00%	93.73%		93.30%	93.10%	93.20%	93.05%
13	94.10%	94.10%	94.10%	94.07%		94.80%	94.30%	94.50%	94.32%		90.90%	88.00%	88.70%	87.97%
14	93.10%	91.90%	92.30%	91.95%		93.90%	93.30%	93.50%	93.31%		92.90%	91.90%	92.20%	91.95%
15	94.20%	94.00%	94.10%	93.98%		94.40%	93.80%	94.00%	93.81%		92.70%	91.60%	91.90%	91.61%

[Halaman ini sengaja dikosongkan]

Lampiran 6

Hasil Uji Statistik *one-way* ANOVA 10 dataset UCI

HASIL UJI DATASET ABALONE

Tests of Normality

METODE		Kolmogorov-Smirnov[a]			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
HASIL UJI DATASET ABALONE	AVERAGE	.326	5	.088	.836	5	.154
	COMPLETE	.251	5	.200*	.868	5	.257
	SINGLE	.326	5	.089	.827	5	.132

*. This is a lower bound of the true significance.
a. Lilliefors Significance Correction

Test of Homogeneity of Variances

Levene Statistic	df1	df2	Sig.
8.703	2	12	.005

Tests of ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	.004	2	.002	6.880	.010
Within Groups	.003	12	.000		
Total	.007	14			

Tests of Duncan[a]

METODE	N	Subset for alpha = 0.05	
		1	2
SINGLE	5	**.9457680**	
AVERAGE	5		.9778940
COMPLETE	5		.9803140
Sig.		1.000	.820

Means for groups in homogeneous subsets are displayed.
a. Uses Harmonic Mean Sample Size = 5.000.

HASIL UJI DATASET BREAST

Tests of Normality

METODE		Kolmogorov-Smirnov[a]			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
HASIL UJI DATASET BREAST	AVERAGE	.378	5	.019	.710	5	.012
	COMPLETE	.234	5	.200*	.918	5	.516
	SINGLE	.353	5	.041	.723	5	.017

*. This is a lower bound of the true significance.
a. Lilliefors Significance Correction

Test of Homogeneity of Variances

Levene Statistic	df1	df2	Sig.
4.648	2	12	.032

Tests of ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	.007	2	.003	.789	.477
Within Groups	.053	12	.004		
Total	.060	14			

Tests of Duncan[a]

METODE	N	Subset for alpha = 0.05
		1
AVERAGE	5	.8864200
SINGLE	5	.8888880
COMPLETE	5	.9333340
Sig.		.309

Means for groups in homogeneous subsets are displayed.
a. Uses Harmonic Mean Sample Size = 5.000.

HASIL UJI DATASET ECOLI

Tests of Normality

METODE		Kolmogorov-Smirnov[a]			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
HASIL UJI DATASET ECOLI	AVERAGE	.342	5	.057	.761	5	.037
	COMPLETE	.237	5	.200*	.961	5	.813
	SINGLE	.389	5	.013	.763	5	.039

*. This is a lower bound of the true significance.
a. Lilliefors Significance Correction

Test of Homogeneity of Variances

Levene Statistic	df1	df2	Sig.
1.595	2	12	.243

Tests of ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	.001	2	.000	2.834	.098
Within Groups	.001	12	.000		
Total	.002	14			

Tests of Duncan[a]

METODE	N	Subset for alpha = 0.05	
		1	2
COMPLETE	5	.9281260	
AVERAGE	5	.9375020	.9375020
SINGLE	5		.9445320
Sig.		.200	.329

Means for groups in homogeneous subsets are displayed.
a. Uses Harmonic Mean Sample Size = 5.000.

HASIL UJI DATASET GLASS

Tests of Normality

METODE		Kolmogorov-Smirnov[a]			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
HASIL UJI DATASET GLASS	AVERAGE	.237	5	.200*	.961	5	.814
	COMPLETE	.217	5	.200*	.936	5	.641
	SINGLE	.241	5	.200*	.821	5	.119

*. This is a lower bound of the true significance.
a. Lilliefors Significance Correction

Test of Homogeneity of Variances

Levene Statistic	df1	df2	Sig.
.049	2	12	.952

Tests of ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	.000	2	.000	.912	.428
Within Groups	.001	12	.000		
Total	.001	14			

Tests of Duncan[a]

METODE	N	Subset for alpha = 0.05
		1
AVERAGE	5	.9836460
SINGLE	5	.9874200
COMPLETE	5	.9893540
Sig.		.230

Means for groups in homogeneous subsets are displayed.
a. Uses Harmonic Mean Sample Size = 5.000.

HASIL UJI DATASET LIBRA

Tests of Normality

METODE		Kolmogorov-Smirnov[a]			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
HASIL UJI DATASET LIBRA	AVERAGE	.254	5	.200*	.914	5	.492
	COMPLETE	.237	5	.200*	.961	5	.814
	SINGLE	.201	5	.200*	.881	5	.314

*. This is a lower bound of the true significance.
a. Lilliefors Significance Correction

Test of Homogeneity of Variances

Levene Statistic	df1	df2	Sig.
.335	2	12	.722

Tests of ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	.000	2	.000	.500	.619
Within Groups	.000	12	.000		
Total	.000	14			

Tests of Duncan[a]

METODE	N	Subset for alpha = 0.05
		1
SINGLE	5	.9875900
AVERAGE	5	.9905100
COMPLETE	5	.9905100
Sig.		.426

Means for groups in homogeneous subsets are displayed.
a. Uses Harmonic Mean Sample Size = 5.000.

HASIL UJI DATASET OCR

Tests of Normality

METODE		Kolmogorov-Smirnov[a]			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
HASIL UJI DATASET OCR	AVERAGE	.221	5	.200*	.953	5	.758
	COMPLETE	.237	5	.200*	.961	5	.814
	SINGLE	.213	5	.200*	.963	5	.826

*. This is a lower bound of the true significance.
a. Lilliefors Significance Correction

Test of Homogeneity of Variances

Levene Statistic	df1	df2	Sig.
.628	2	12	.550

Tests of ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	.000	2	.000	.738	.499
Within Groups	.000	12	.000		
Total	.000	14			

Tests of Duncan[a]

METODE	N	Subset for alpha = 0.05
		1
SINGLE	5	.9991420
AVERAGE	5	.9992080
COMPLETE	5	.9995380
Sig.		.302

Means for groups in homogeneous subsets are displayed.
a. Uses Harmonic Mean Sample Size = 5.000.

HASIL UJI DATASET ROBOT

Tests of Normality

METODE		Kolmogorov-Smirnov[a]			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
HASIL UJI DATASET ROBOT	AVERAGE	.237	5	.200*	.961	5	.814
	COMPLETE	.286	5	.200*	.913	5	.486
	SINGLE	.136	5	.200*	.987	5	.967

*. This is a lower bound of the true significance.
a. Lilliefors Significance Correction

Test of Homogeneity of Variances

Levene Statistic	df1	df2	Sig.
.613	2	12	.558

Tests of ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	.000	2	.000	5.652	.019
Within Groups	.000	12	.000		
Total	.000	14			

Tests of Duncan[a]

METODE	N	Subset for alpha = 0.05	
		1	2
SINGLE	5	**.9981700**	
AVERAGE	5		.9987680
COMPLETE	5		.9988120
Sig.		1.000	.840

Means for groups in homogeneous subsets are displayed.
a. Uses Harmonic Mean Sample Size = 5.000.

HASIL UJI DATASET SATIMAGE

Tests of Normality[a,b]

METODE	Kolmogorov-Smirnov[c]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
HASIL UJI DATASET SATIMAGE SINGLE	.242	5	.200*	.821	5	.118

*. This is a lower bound of the true significance.
a. HASIL UJI DATASET SATIMAGE is constant when METODE = AVERAGE. It has been omitted.
b. HASIL UJI DATASET SATIMAGE is constant when METODE = COMPLETE. It has been omitted.
c. Lilliefors Significance Correction

Test of Homogeneity of Variances

Levene Statistic	df1	df2	Sig.
16.297	2	12	.000

Tests of ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	.000	2	.000	19.927	.000
Within Groups	.000	12	.000		
Total	.000	14			

Tests of Duncan[a]

METODE	N	Subset for alpha = 0.05	
		1	2
AVERAGE	5	.9989100	
SINGLE	5	.9989140	
COMPLETE	5		1.0000000
Sig.		.984	1.000

Means for groups in homogeneous subsets are displayed.
a. Uses Harmonic Mean Sample Size = 5.000.

HASIL UJI DATASET WINE

Tests of Normality

METODE		Kolmogorov-Smirnov[a]			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
HASIL UJI DATASET WINE	AVERAGE	.367	5	.026	.684	5	.006
	COMPLETE	.300	5	.161	.883	5	.325
	SINGLE	.473	5	.001	.552	5	.000

a. Lilliefors Significance Correction

Test of Homogeneity of Variances

Levene Statistic	df1	df2	Sig.
5.315	2	12	.022

Tests of ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	.000	2	.000	1.129	.355
Within Groups	.002	12	.000		
Total	.002	14			

Tests of Duncan[a]

METODE	N	Subset for alpha = 0.05
		1
AVERAGE	5	.9806000
SINGLE	5	.9865700
COMPLETE	5	.9925400
Sig.		.178

Means for groups in homogeneous subsets are displayed.
a. Uses Harmonic Mean Sample Size = 5.000.

HASIL UJI DATASET YEAST

Tests of Normality

METODE		Kolmogorov-Smirnov[a]			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
HASIL UJI DATASET YEAST	AVERAGE	.233	5	.200*	.852	5	.202
	COMPLETE	.213	5	.200*	.952	5	.751
	SINGLE	.377	5	.019	.811	5	.099

*. This is a lower bound of the true significance.
a. Lilliefors Significance Correction

Test of Homogeneity of Variances

Levene Statistic	df1	df2	Sig.
2.322	2	12	.140

Tests of ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	.002	2	.001	4.893	.028
Within Groups	.002	12	.000		
Total	.004	14			

Tests of Duncan[a]

METODE	N	Subset for alpha = 0.05	
		1	2
SINGLE	5	.9123720	
AVERAGE	5	.9300000	.9300000
COMPLETE	5		.9372900
Sig.		.052	.391

Means for groups in homogeneous subsets are displayed.
a. Uses Harmonic Mean Sample Size = 5.000.

BIODATA PENULIS

Meida Cahyo Untoro, lahir di Kediri 18 Mei 1989. Penulis merupakan anak kedua dari pasangan Maryanto dan Karsini. Penulis juga merupakan penikmat musik, gemar berolahraga, keterkaitan dengan elektro, dan otomotif kendaraan (modifikasi dan bengkel).

Penulis menempuh pendidikan formal di SDN Sambirobyong II (1995 – 2001), SMP Negeri 1 Pagu (2001 – 2004), SMA Negeri 1 Papar (2004 – 2007), S-1 Sistem Komputer STIKOM BALI dengan bidang minat *User Interface* (2009-2012). Pada tahun 2016 penulis melanjutkan pendidikan Magister di Institut Teknologi Sepuluh Nopember dengan jurusan Teknik Informatika. Saat menempuh pendidikan S2 penulis sebagai Tenaga Pendidik di Institut Ilmu Kesehatan Bhakti Wiyata dengan Program Studi Rekam Medis Informasi Kesehatan.

Penulis menaruh minat pada topik Dasar dan Terapan Komputasi (DTK). Penulis dapat dihubungi via e-mail dengan alamat meidacahyountoro@gmail.com