



Artikel Penelitian

## Penentuan Jurusan Siswa SMA Berdasarkan Tes Minat Bakat Menggunakan Metode *Single, Complete* dan *Average Linkage*

Alvian M. Sroyer<sup>a</sup>, Habel Saud<sup>b</sup> dan Felix Reba<sup>c,\*</sup>

<sup>a</sup>Program Studi Matematika, Fakultas MIPA, Universitas Cenderawasih, Papua, Indonesia

<sup>b</sup>Program Studi Bimbingan dan Konseling, FKIP, Universitas Cenderawasih, Papua, Indonesia

### INFORMASI ARTIKEL

#### Sejarah Artikel:

Diterima Redaksi: 09 Juni 2022

Revisi Akhir: 14 Agustus 2022

Diterbitkan Online: 31 Agustus 2022

### KATA KUNCI

Tes Minat Bakat,

Jurusan Siswa SMA,

Pengelompokan,

Metode Cluster

### KORESPONDENSI

E-mail: [felix.reba85@gmail.com](mailto:felix.reba85@gmail.com)\*

### A B S T R A C T

Menurut data yang dirilis oleh United Nations Children's Fund (Unicef) tentang situasi pendidikan di Papua, dimana 30% siswa di Provinsi Papua secara umum tidak menyelesaikan Pendidikan SD dan SMP. Di wilayah pedalaman Papua, kira-kira 50% peserta didik jenjang SD dan 73% SMP lebih memilih putus sekolah. Jika fenomena ini dibiarkan berlanjut, maka para siswa akan kesulitan dalam menentukan jurusan yang sesuai dengan kemampuan mereka baik di jenjang SMA, maupun jika para siswa ingin melanjutkan ke perguruan tinggi. Beberapa tahun terakhir ini, beberapa SMA di kota Jayapura sudah menggunakan hasil Intelligenz Struktur Test (IST) dalam menentukan jurusan siswa. Tim tes psikologi yang sering melakukan tes ini adalah para dosen dari Program Studi Bimbingan dan Konseling (BK) UNCEN. Namun tim Prodi BK juga terkadang mengalami kendala, jika siswa yang mengikuti tes dalam jumlah yang banyak. Akibatnya, waktu yang diperlukan untuk menyelesaikan satu subtes IST relatif lama. Terkadang terjadi error yang tak terduga dalam menentukan hasil tes menggunakan norma. Sehingga penelitian ini, akan digunakan tiga metode cluster yaitu : *Single Linkage*, *Complete Linkage* dan *Average Linkage* dalam menentukan jurusan siswa SMA. Penggunaan tiga metode ini tujuannya adalah bagaimana mengetahui metode yang paling sesuai dalam menentukan jurusan siswa SMA. Hasil dendogram memperlihatkan bahwa, jumlah jurusan yang dapat dibentuk adalah dua jurusan. Selanjutnya hasil analisis menunjukkan bahwa, dari 305 siswa kelas X yang mengikuti tes minat dan bakat, terdapat hanya 7 siswa yang memenuhi syarat untuk masuk pada jurusan IPA dan 298 siswa memenuhi syarat untuk masuk pada jurusan IPS. Harapannya metode terbaik dapat menjadi rujukan sebagai bahan validasi. Selain itu, tim tes psikologi Prodi BK UNCEN dapat menjangkau seluruh SMA dan Sederajat di Provinsi Papua untuk melakukan tes psikologi dalam penentuan jurusan siswa.

## 1. PENDAHULUAN

Pendidikan memiliki peranan yang sangat penting untuk meningkatkan Sumber Daya Manusia (SDM), karena Pendidikan mempunyai pengaruh yang signifikan terhadap produktivitas dan juga kualitas hidup masyarakat. [1]. Hingga saat ini Pendidikan di Indonesia masih menjadi perhatian khusus, karena kualitasnya masih sangat memprihatinkan. Di Kawasan timur Indonesia kualitas pendidikan masih tergolong rendah, salah satunya di provinsi Papua. Para siswa di provinsi Papua masih kesulitan untuk mendapatkan Pendidikan yang layak, bahkan pendidikan dasar sekalipun [2]. Berdasarkan data yang dirilis oleh Badan

Pusat (BPS) Statistik provinsi Papua terkait kurangnya akses pendidikan yang diukur dengan angka Indeks Pembangunan Manusia (IPM), Papua masih berada di urutan terbawah dengan IPM 59.09 [32]. Menurut data yang dirilis oleh United Nations Children's Fund (Unicef) tentang situasi pendidikan di Papua, dimana 30% siswa di Provinsi Papua secara umum tidak menyelesaikan Pendidikan SD dan SMP. Di wilayah pedalaman Papua, kira-kira 50% peserta didik jenjang SD dan 73% SMP lebih memilih putus sekolah [3]. Jika fenomena ini dibiarkan berlanjut, maka akan berpengaruh terhadap pendidikan siswa di jenjang Sekolah Menengah Atas (SMA). Dimana para siswa akan kesulitan dalam menentukan jurusan yang sesuai dengan kemampuan mereka. Contohnya jika siswa ingin memutuskan

untuk masuk pada jurusan IPA atau IPS dibutuhkan beberapa kemampuan khusus terkait jurusan yang dipilih.

Pemilihan jurusan juga tidak boleh dipandang sepele, karena akan sangat berdampak terhadap kegiatan akademik di jenjang selanjutnya jika para siswa ingin melanjutkan ke perguruan tinggi, juga mempengaruhi pemilihan bidang ilmu yang sesuai [4]. Sebenarnya penentuan jurusan siswa SMA telah ditetapkan dalam kurikulum 2013, yang mana proses penjurusan sudah dapat dimulai dari kelas X. Dimana proses penentuan jurusan dapat menggunakan nilai akademik siswa, namun hal ini tidak secara langsung menjadi tolak ukur karena dapat juga menimbulkan ketidaksesuaian minat dan bakat. Penentuan jurusan juga dapat menggunakan angket minat dan bakat yang dilakukan oleh tim bimbingan dan konseling [5]. Dalam beberapa tahun terakhir ini beberapa SMA di kota Jayapura sudah menggunakan hasil *Intelligenz Struktur Test (IST)* dalam menentukan jurusan siswa. Tim tes psikologi yang sering melakukan tes ini adalah, para dosen dari Program Studi Bimbingan dan Konseling (BK) UNCEN. Hasilnya beberapa sekolah merasa terbantu dengan adanya program ini. Cara menentukan jurusan yang selama ini digunakan adalah, hasil IST yang diperoleh kemudian digunakan beberapa norma-norma baku yang telah ditentukan. Namun tim Prodi BK juga terkadang mengalami kendala jika siswa yang mengikuti tes dalam jumlah yang banyak. Akibatnya waktu yang diperlukan untuk menyelesaikan satu paket IST relatif lama bahkan dapat memakan waktu hingga 2-3 minggu. Sehingga terkadang terjadi error yang tak terduga dalam menentukan hasil tes dengan norma. Sebenarnya ada metode dibidang data mining yang dapat digunakan untuk penentuan jurusan berdasarkan hasil tes minat bakat seperti metode clustering. Dimana metode clustering merupakan metode yang sangat sederhana, tingkat akurasi yang sangat tinggi dan membutuhkan waktu yang relatif lebih singkat. Oleh karena itu, sudah waktunya Prodi BK UNCEN perlu adanya kolaborasi dengan bidang-bidang lain untuk menyelesaikan masalah yang dihadapi seperti bidang ilmu data.

Sebenarnya sudah banyak peneliti yang menggunakan metode *cluster* dalam penentuan jurusan siswa SMA seperti [6], [7], [8], [9], namun data yang digunakan adalah data nilai mata pelajaran. Tidak banyak yang menggunakan metode *clustering* menggunakan data tes minat dan bakat yang dilakukan oleh tim BK seperti [10] dalam penelitiannya mereka menggunakan metode *k-Nearest Neighbor (k-NN)*. Sehingga pada penelitian ini, akan digunakan data tes minat dan bakat yang dilakukan oleh tim BK UNCEN dengan tiga metode cluster yaitu : *Single Linkage*, *Complete Linkage* dan *Average Linkage* dalam menentukan jurusan siswa SMA. Penggunaan tiga metode cluster pada data tes minat bakat tidak hanya untuk membandingkan kebaikan suatu metode, karena setiap metode memiliki kelebihan dan kekurangan tergantung pada kasus yang diteliti. Namun penggunaan tiga metode tujuannya adalah bagaimana mengetahui metode yang paling sesuai dalam menentukan jurusan siswa SMA. Harapannya metode terbaik dapat menjadi rujukan untuk membantu tim tes psikologi Prodi BK UNCEN dalam menganalisis hasil IST siswa SMA. Sehingga metode terbaik bukan satu-satunya metode yang harus digunakan, namun dapat digunakan sebagai bahan validasi. Bagaimana pun juga norma yang selama ini digunakan merupakan alat yang lebih valid, karena terdapat beberapa parameter yang tidak dapat diselesaikan

berdasarkan data dan komputasi yaitu yang berkaitan dengan pengalaman seseorang.

## 2. METODE

Pada bagian metode penelitian ini, akan dijelaskan langkah-langkah menggunakan metode cluster dalam menentukan jurusan siswa SMA yang mungkin dengan data yang digunakan. Awalnya akan dimulai dari penjelasan terkait Jurusan, Tes Intelegensi, Syarat untuk masuk jurusan IPA atau IPS, Standarisasi data, Menggambar Dendogram dan Scatter Plot dari Data, Menentukan Nilai Koefisien Korelasi dari ketiga metode, Menganalisis data berdasarkan metode terbaik. Berikut adalah penjelasan setiap Langkah :

### 2.1 Penjurusan

Peminatan siswa SMA dan Sederajat merupakan upaya mempersiapkan manusia yang dapat hidup mandiri, dewasa dan berkompeten di masyarakat. Kelompok minat akademik di setiap jenjang pendidikan SMA dan sederajat umumnya terbagi menjadi tiga kelompok, yang pertama cluster bidang minat sains seperti Matematika dan Ilmu Pengetahuan Alam (IPA), kedua kelompok bidang minat Ilmu Pengetahuan Sosial (IPS), dan ketiga kelompok yang minat Bahasa dan Seni. Beberapa aspek yang sering digunakan dalam memilih dan menetapkan bidang minat peserta didik adalah dengan melakukan tes potensi minat dan bakat. Tes ini biasanya dilakukan oleh tim tes BK. Data hasil tes kemudian dianalisis menggunakan beberapa norma untuk mendeteksi potensi siswa. Hasil analisis ini yang kemudian dapat digunakan sebagai rekomendasi bagi sekolah untuk menentukan bidang minat para siswa [5].

### 2.2 Tes Intelegensi

*Intelligenz Struktur Test (IST)* merupakan Tes intelegensi yang sering digunakan guna mengetahui potensi kecerdasan kognitif peserta didik. IST awalnya dikembangkan oleh Rudolf Amthauer yang berkebangsaan Jerman yaitu sekitar tahun 1953. Alat ini telah dirancang dengan sangat baik sehingga dapat menggambarkan komponen kecerdasan yang terdiri dari kecerdasan verbal, figural, numerik dan umum. IST sendiri terdiri dari 9 bagian, yang berisi kurang lebih 176 soal. Berikut adalah sembilan subtes [11] :

- a. *Analogien (AN)* – Mencari hubungan kata
- b. *Form Auswahl (FA)* – Menyusun Bentuk
- c. *Gemeinsamkeiten (GE)* – Mencari Kata yang Mengandung Dua Pengertian
- d. *Merk Aufgaben (ME)* – Mengingat Kata
- e. *Rechen Aufgaben (RA)* – Hitungan Sederhana
- f. *Satzerganzung (SE)* – Melengkapi kalimat
- g. *Wurfal Aufgaben (WU)* – Kubus
- h. *Wortauswahl (WA)* – Mencari kata yang berbeda
- i. *Zahlen Reihen (ZR)* – Deret angka

### 2.3 Angka IST penjurusan IPA

Angka IST adalah angka yang dipakai untuk menentukan siswa bisa masuk di suatu jurusan. Contohnya, jika seorang siswa

memiliki nilai RA, ZR, FA dan WU lebih tinggi dibandingkan dengan yang lain, maka siswa tersebut dapat direkomendasikan untuk masuk pada jurusan IPA. Nilai ini sangat membantu karena pada saat siswa tersebut ingin memilih masuk pada jurusan IPA, yang akan diutamakan yaitu pelajaran seperti : matematika, fisika, kimia dan biologi [12].

#### 2.4 Angka IST penjurusan IPS

Sama halnya dengan nilai IST untuk penjurusan IPA. Agar seorang siswa dapat masuk pada jurusan IPS, maka siswa tersebut memerlukan nilai SE, WA, GE dan ME yang harus lebih tinggi dibandingkan dengan yang lain. Sebab pada saat siswa masuk pada jurusan IPS, mata pelajaran yang akan ditemui adalah : ekonomi, geografi dan sosiologi [12].

#### 2.5 Standarisasi Data

Tujuan data perlu distandarisasi adalah, jika pada variabel yang akan diteliti terdapat perbedaan ukuran satuan atau tiap item data memiliki rasio yang cukup besar. Perbedaan rasio pada item data yang cukup besar dapat menyebabkan perhitungan jarak dan analisis cluster tidak valid atau memiliki standar deviasi yang besar. Sehingga perlu dilakukan standarisasi terhadap data aktual dengan cara transformasi menggunakan uji *z-score* sebelum melakukan analisis yang lebih lanjut [13]. Formula untuk metode transformasi data aktual seperti yang ditunjukkan pada persamaan 1 dibawah ini :

$$z_i = \frac{x_i - \bar{x}}{s}, \quad \text{untuk } i = 1, 2, 3, \dots, n \quad (1)$$

Dimana :

- $x_i$  : Data Ke- $i$
- $\bar{x}$  : Rata-rata Data Aktual
- $s$  : Standar Deviasi

Dengan  $s$  adalah standar deviasi yang dirumuskan di bawah ini :

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x}_i)^2}{n-1}}, \quad (2)$$

#### 2.6 Dendrogram

Ada dua cara untuk mengelompokkan data, hierarkis dan non-hierarkis. Namun pada penelitian ini, penulis hanya memaparkan metode hierarki. Metode hierarki adalah agregasi (penggabungan) atau splitting (pembelahan). Untuk metode aglomerasi sendiri, setiap penelitian pada awalnya diasumsikan merupakan cluster tersendiri dimana terdapat cluster yang sama dengan jumlah penelitiannya. Selain itu, jika terdapat dua cluster yang paling mirip maka keduanya akan digabungkan menjadi satu cluster baru, mengurangi jumlah cluster sebanyak satu per level. Sementara itu, solusi akan dimulai dengan cluster besar yang berisi semua pengamatan. Selain itu, jika pengamatan tidak mirip, mereka proses selanjutnya adalah dipisahkan untuk membentuk kelompok yang lebih kecil. Proses ini berlanjut hingga tiap pengamatan terbentuk dalam kelompoknya sendiri [14] dan [15]. Proses pengelompokan hierarkis dimulai dengan dua atau lebih objek yang paling mirip. Selain itu, perhitungan diteruskan ke objek terdekat atau terdekat kedua. Terdapat hierarki yang jelas antara tiap objek, mulai dari yang paling mirip sampai yang paling tidak mirip, sehingga cluster membentuk dendrogram atau semacam diagram pohon [16].

#### 2.7 Koefisien Korelasi

Koefisien korelasi memiliki dua nilai yang dipakai sebagai ukuran ada tidaknya hubungan linear. Yang pertama, jika koefisien korelasi bernilai 1, maka terdapat hubungan linier yang positif. Yang kedua, jika nilai koefisien korelasi bernilai -1, maka terdapat hubungan linier yang negatif. Yang ketiga, jika nilai korelasi terletak di antara -1 dan +1, maka nilai ini menunjukkan tingkat dependensi linier antara dua variabel. Artinya, semakin dekat dengan -1 atau +1, semakin kuat korelasi antara kedua variabel tersebut. [17], [18] dan [18]. Dengan koefisien korelasi yang dirumuskan sebagai berikut :

$$R = \frac{\sum_{i=1}^n (x_1 - \bar{x}_1)(x_2 - \bar{x}_2)}{\sqrt{\sum_{i=1}^n (x_1 - \bar{x}_1)^2} \sqrt{\sum_{i=1}^n (x_2 - \bar{x}_2)^2}}, \quad i = 1, 2, \dots, n \quad (3)$$

Jika variabel tersebut independen, maka nilai korelasi sama dengan 0. Namun, koefisien korelasi hanya mendeteksi ketergantungan linier antara dua variabel, sehingga tidak berlaku sebaliknya. Jika asumsi multikolinearitas terpenuhi, mak jarak Euclidean boleh digunakan dalam menentukan kesamaan [19].

#### 2.8 Analisis Cluster

Analisis cluster merupakan teknik analisis statistika yang digunakan dalam mengklasifikasikan kumpulan objek menjadi dua atau lebih kelompok sesuai kesamaan karakteristik objek tersebut. Analisis ini adalah teknik analisis yang tujuannya untuk menyeleksi objek dari beberapa kelompok dimana satu kelompok memiliki sifat yang berbeda dengan yang lain. Pada analisis cluster, setiap kelompok homogen di antara anggota kelompok, atau variasi objek dalam kelompok yang terbentuk sekecil mungkin. Objek-objek di setiap cluster mirip satu sama lain dan cenderung sangat berbeda (tidak sama) dengan objek-objek di cluster lain. Pengelompokan dilakukan berdasarkan kemiripan antar objek. Kemiripan dapat diperoleh dengan cara meminimalkan jarak antara objek di dalam cluster dan memaksimalkan jarak di antara cluster. Secara keseluruhan solusi cluster bergantung pada setiap variabel yang digunakan sebagai dasar dalam menilai suatu kemiripan. Pengurangan dan penambahan variabel yang terkait dapat dikatakan sebagai suatu proses dalam menganalisa baik atau tidaknya suatu proses pembentukan cluster. Berikut adalah beberapa metode jarak yang sering digunakan dalam metode pengelompokan [20] dan [21] :

##### 2.8.1 Euclidean Distance

Dalam matematika, jarak Euclidean merupakan panjang segmen garis antara dua titik. Titik ini dapat dihitung dari koordinat Cartesius dari tiap titik menggunakan teorema Pythagoras, oleh karenanya sering disebut jarak Pythagoras. Metode penghitungan jarak meliputi bidang Euclidean dua dimensi, tiga dimensi hingga dimensi ke  $n$ . Sehingga untuk mengukur kemiripan data menggunakan rumus jarak Euclidean, gunakan rumus berikut [22]:

$$d(x, y) = |x - y| = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (4)$$

Dimana :

$x_i$  = Data Pada Pusat Cluster ke  $i$

$y_i$  = Data Pada Setiap Data ke  $i$

$y$  = Data pada Atribut  
 $x$  = Data pada Pusat Cluster  
 $i$  = Setiap Data  
 $n$  = Jumlah Data  
 $d$  = Jarak Antara  $x$  dan  $y$

### 2.8.2 Metode Single Linkage

Metode tautan tunggal adalah salah satu dari beberapa metode pengelompokan hierarkis. Metode ini didasarkan pada agregasi cluster, di mana setiap langkah menggabungkan dua cluster yang berisi pasangan elemen terdekat yang belum terdapat dalam cluster yang sama. Prosedur ini biasanya dimulai dengan menentukan objek dengan nilai jarak terkecil dalam matriks proximity  $D = \{d_{ik}\}$ . Objek-objek yang memiliki jarak terkecil disatukan dalam satu kelompok, misalkan  $U$  dan  $V$  adalah dua objek yang akan dilakukan pengelompokan, maka akan diperoleh kelompok ( $UV$ ). Untuk mencari jarak antara grup ( $UV$ ) dan grup  $W$  atau grup lainnya, gunakan rumus berikut [23], [24] dan [25]:

$$d_{(UV)W} = \min\{d_{UW}, d_{VW}\} \quad (5)$$

Dimana :

$d_{UW}$  : Jarak Terdekat Dari Kelompok  $U$  dan  $W$

$d_{VW}$  : Jarak Terdekat Dari Kelompok  $V$  dan  $W$

### 2.8.3. Metode Complete Linkage

Dalam metode pautan lengkap (complete linkage) clustering didasarkan pada jarak maksimum antara satu objek dengan objek lainnya. Dengan metode ini, perlu dicari jarak terpendek terlebih dahulu  $D = \{d_{ik}\}$  dan gabungkan objek yang sesuai seperti  $U$  dan  $V$  untuk mendapatkan cluster ( $UV$ ). Untuk langkah selanjutnya dari algoritma ini, jarak antara ( $UV$ ) dan cluster  $W$  lainnya dihitung sebagai berikut [26], dan [27]:

$$d_{(UV)W} = \max\{d_{UW}, d_{VW}\} \quad (6)$$

Besaran-besaran  $d_{UW}$  dan  $d_{VW}$  berturut-turut adalah jarak terjauh antara cluster-cluster  $U$  dan  $W$  dan juga klater-cluster  $V$  dan  $W$ .

### 2.8.4 Metode Average Linkage

Pada metode average linkage atau jarak rata-rata antar observasi, pengelompokan cluster dimulai dari tengah atau pasangan observasi dengan jarak paling mendekati jarak rata-rata. Prosedur ini dilakukan dengan meminimumkan jarak rata-rata semua pasangan objek yang berasal dari satu kelompok terhadap kelompok lainnya. Jadi, awalnya perlu ditemukan jarak terpendek dalam  $D = \{d_{ik}\}$  dan menggabungkan objek-objek yang selaras. Misalnya  $U$  dan  $V$  untuk mendapat cluster ( $UV$ ). Untuk langkah (iii) dari algoritma di atas jarak-jarak antara ( $UV$ ) dan cluster  $W$  yang lain dihitung dengan cara [28], [29], [30] dan [31]:

$$d_{(UV)W} = \frac{\sum_i \sum_k d_{ik}}{N_{UV} N_W} \quad (7)$$

Dimana  $d_{ik}$  merupakan jarak antara objek di dalam cluster ( $UV$ ) dan objek ke dalam cluster  $W$ , dan  $N_{UV}$  dan  $N_W$  adalah banyaknya item-item dalam cluster ( $UV$ ) dan  $W$ .

Pengukuran jarak memainkan peran yang sangat penting dalam menentukan kesamaan atau keteraturan antara data dan item. Hal ini dilakukan untuk mengetahui bagaimana data-data tersebut

saling terhubung, serupa atau tidak, dan teknik pengukuran jarak yang diperlukan untuk membandingkannya. Dalam proses clustering, fase (aksesibilitas) yang menentukan atau menggambarkan nilai kuantitatif kesamaan dan ketidaksamaan data memegang peranan yang sangat penting. Metode cluster banyak digunakan untuk jarak Euclidean, jarak Manhattan, jarak Minkowski, dll.

## 3. HASIL DAN PEMBAHASAN

Pada bagian hasil dan pembahasan, akan ditunjukkan jumlah jurusan yang dapat terbentuk. Penentuan jurusan ini menggunakan salah satu metode terbaik dari ketiga metode cluster dengan mengikuti langkah-langkah yang telah dijelaskan pada dasar teori :

### 3.1 Jenis dan Sumber Data

Data yang digunakan dalam penelitian ini adalah data hasil tes IST minat dan bakat siswa SMA N 4 Jayapura, yang lakukan oleh tim tes psikologi BK dari Prodi BK UNCEN. Dimana peneliti merupakan salah satu dari tim BK yang memberikan tes IST, sehingga data yang digunakan merupakan data primer. Pada saat dilakukan tes, terdapat 13 subtes IST yang diberikan kepada siswa. Namun pada penelitian ini, hanya digunakan 10 subtes dengan jumlah siswa 305, karena dianggap 10 item ini yang relevan dengan tujuan penelitian. Pada tabel 1 dan 2 di bawah ini, disajikan data hasil penelitian:

Tabel 1. Data Tes Bakat Siswa SMA N 4 Jayapura

No	RPM		Bakat			
	IQ	SE	ZR	WA	ME	FA
1	106	15	5	3	20	75
2	62	3	1	0	5	85
3	94	0	0	0	5	30
4	91	1	10	1	35	90
5	116	1	25	5	40	75
...	...	...	...	...	...	...
303	120	10	10	5	65	45
304	97	0	10	0	5	15
305	93	0	0	0	10	1

Tabel 2. Data Tes Minat Siswa SMA N 4 Jayapura

No	Minat			
	AN	GE	WU	RA
1	80	50	60	70
2	30	40	60	50
3	90	50	70	30
4	80	80	80	40
5	99	5	20	40
...	...	...	...	...
303	60	30	20	95
304	95	50	30	10
305	20	1	1	3

### 3.2 Standarisasi Data

Sebelum melakukan standarisasi data minat dan bakat siswa SMA N 4 Jayapura, perlu dicek kelengkapan data masing-masing

siswa. Jika data sudah sesuai, maka data harus disimpan kedalam beberapa format sesuai kebutuhan software. Pada penelitian ini software yang digunakan untuk analisis data adalah Matlab R2018b, maka data excel akan disimpan dalam format csv. Selanjutnya data diimport ke Matlab untuk dianalisis. Tujuan dari standarisasi data adalah, agar data dari setiap subtes IST memiliki rasio yang sama. Selanjutnya data distandarisasi menggunakan persamaan 1 di atas. Sehingga tabel 3 di bawah ini adalah hasil standarisasi data dengan Matlab R2018b.

Tabel 3. Hasil Standarisasi Data Bakat

No	RPM		Bakat			
	IQ	SE	ZR	WA	ME	FA
1	0.49	1.66	-0.38	-0.07	-0.29	1.04
2	-2.44	-0.16	-0.65	-0.44	-0.96	1.35
3	-0.31	-0.61	-0.71	-0.44	-0.96	-0.37
4	-0.51	-0.46	-0.05	-0.31	0.39	1.51
5	1.16	-0.46	0.94	0.18	0.61	1.04
...	...	...	...	...	...	...
303	1.43	0.90	-0.05	0.18	1.74	0.10
304	-0.11	-0.61	-0.05	-0.44	-0.96	-0.84
305	-0.38	-0.61	-0.71	-0.44	-0.74	-1.28

Tabel 4. Hasil Standarisasi Data Minat

No	Minat			
	AN	GE	WU	RA
1	0.62	0.35	0.74	0.68
2	-0.93	-0.01	0.74	0.05
3	0.93	0.34	1.15	-0.58
4	0.62	1.41	1.57	-0.26
5	1.21	-1.25	-0.93	-0.26
...	...	...	...	...
303	0.00	-0.36	-0.93	1.47
304	1.09	0.34	-0.51	-1.21
305	-1.24	-1.39	-1.72	-1.43

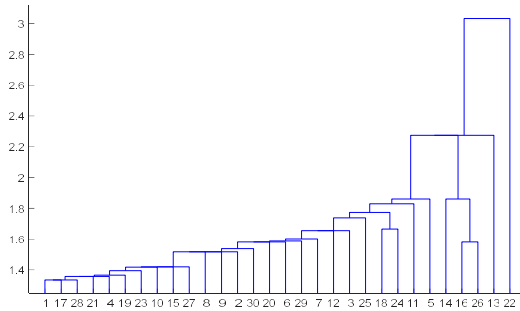
Setelah data distandarisasi, selanjutnya dibuat dendrogram untuk data siswa menggunakan ketiga metode cluster.

### 3.3 Dendrogram Tiga Metode

Setelah data distandarisasi, selanjutnya digambarkan dendrogram data untuk setiap metode. Salah satu tujuan menggambarkan dendrogram adalah, agar dapat diketahui cluster maksimal dari data. Berikut adalah gambar dendrogram untuk ketiga metode menggunakan Matlab R2018b :

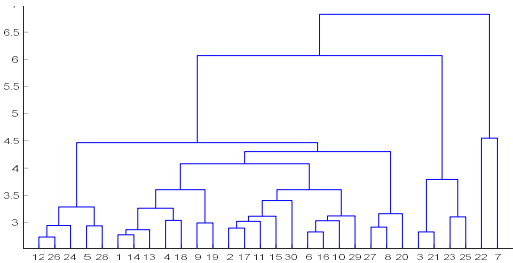
#### 1. Single Linkage

#### 1. Metode Single



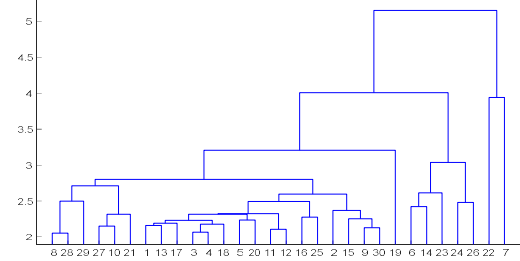
Gambar 1. Dendrogram Single Linkage

#### 2. Metode Complete



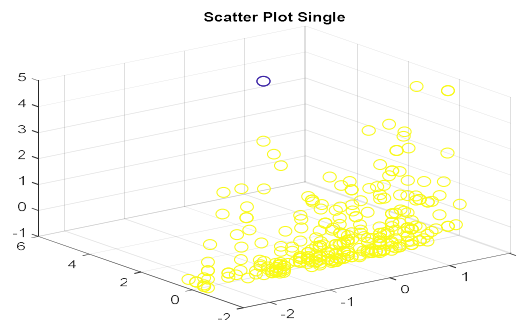
Gambar 2. Dendrogram Complete Linkage

#### 3. Metode Average



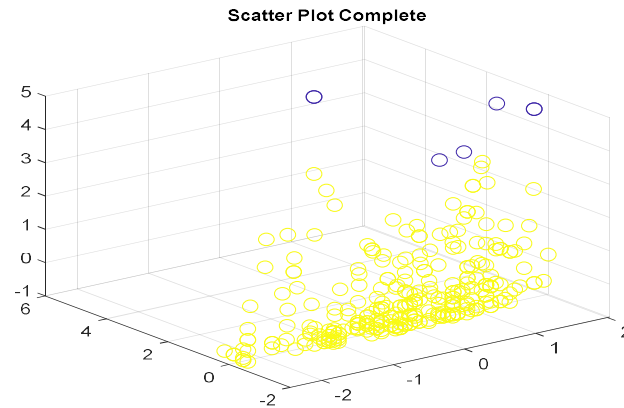
Gambar 3. Dendrogram Average Linkage

Berdasarkan model dendrogram pada gambar 1, 2 dan 3 terlihat jelas bahwa, jumlah jurusan yang mungkin terbentuk dari data siswa adalah dua jurusan. Selain dendrogram, diberikan juga gambar scatter plot dari data untuk masing-masing metode seperti gambar 4, 5, dan 6. Berikut adalah scatter plot menggunakan Matlab R2018b:



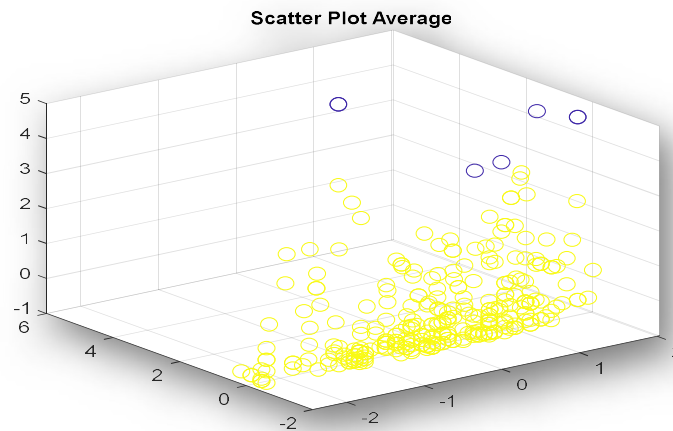
Gambar 4. Scatter Plot Single Linkage

## 2. Complete Linkage



Gambar 5. Scatter Plot Complete Linkage

## 3. Average Linkage



Gambar 6. Scatter Plot Average Linkage

Berbeda dengan gambar 4, scatter plot pada gambar 5 dan 6 terlihat bahwa metode *complete* dan *average linkage* membagi data mahasiswa menjadi dua jurusan dengan jumlah kelompok satu dan kelompok dua yang sama. Sehingga akan digunakan koefisien korelasi untuk menentukan metode terbaik yang akan digunakan untuk menentukan jurusan siswa. Penggunaan scatter plot dalam penelitian ini

### 3.4 Koefisien Korelasi

Berbeda dengan penelitian lainnya, dimana koefisien korelasi biasanya diletakan di bagian akhir hasil dan pembahasan, pada penelitian ini nilai koefisien korelasi justru ditentukan terlebih dahulu. Tujuan menghitung koefisien korelasi adalah, agar dapat diketahui metode terbaik dari ketiga metode untuk mengelompokkan data siswa. Untuk mencari koefisien korelasi, yang pertama dilakukan adalah menentukan jarak euclidean antara pasangan pengamatan dari data siswa. Yang kedua adalah menentukan jarak kofenetik antar objek. Yang ketiga menentukan jarak koefisien korelasi kofenetik antar objek menggunakan persamaan 3. Berikut adalah koefisien korelasi tiga metode menggunakan Matlab R2018b, maka hasilnya adalah sebagai berikut :

Tabel 5. Nilai Koefisien Korelasi Tiga Metode Cluster

Koefisien Korelasi		
Single linkage	Complete linkage	Average linkage
0.5697	0.5069	0.6718

Berdasarkan tabel 5 di atas, terlihat bahwa metode *average linkage* merupakan metode terbaik. Karena memiliki nilai koefisien korelasi 0.6718 yang lebih dekat dengan +1 dibandingkan dengan *single* dan *complete linkage*. Sehingga metode *average linkage* yang akan digunakan untuk membagi 305 siswa ke dalam dua jurusan.

### 3.5 Metode Average Linkage

Setelah diketahui metode terbaik adalah metode *average linkage* dan banyak jurusan yang dapat dibentuk adalah dua jurusan, maka dilanjutkan analisis terperinci menggunakan persamaan 7. Langkah pertama adalah, data yang telah dinormalisasi diubah ke dalam matriks jarak menggunakan persamaan 4. Langkah kedua adalah, menentukan nilai rata-rata dari setiap *cluster* yang terbentuk dari setiap subtes IST. Langkah ketiga adalah, menentukan jumlah siswa untuk setiap jurusan diikuti dengan nilai rata-rata subtes tiap *cluster*. Seperti yang telah dibahas pada

dasar teori bahwa yang masuk jurusan IPA harus memiliki nilai RA, ZR, FA dan WU yang lebih tinggi dibandingkan yang lain, maka hasil analisis disajikan pada tabel 6 di bawah ini :

Tabel 6. Pengelompokan Jurusan IPA dengan *Average Linkage*

Jmlh Siswa	IQ	ZR	FA	WU	RA	Jurusan
7	118.86	70.71	63.57	38.57	80.71	IPA
298	98.16	9.35	41.38	42.54	47.53	IPS

Tujuan dimasukkannya rata-rata nilai jurusan IPS Pada tabel 6 adalah, agar terlihat bahwa rata nilai RA, ZR, FA dan WU dari 298 siswa lebih rendah dari 7 siswa yang masuk pada jurusan IPA.

Selanjutnya, seperti yang telah dibahas pada dasar teori bahwa yang masuk jurusan IPS harus memiliki nilai SE, WA, GE dan ME yang lebih tinggi dibandingkan yang lain, maka hasil analisis disajikan pada tabel 7 di bawah ini :

Tabel 7. Pengelompokan Minat dengan Metode *Average Linkage*

Jmlh Siswa	SE	WA	ME	AN	GE	Jurusan
298	16.43	43.57	66.43	65.71	21.43	IPS
7	3.76	2.58	25.52	59.77	40.72	IPA

Sama seperti tabel 6, pada tabel 7 juga dilakukan hal yang serupa yaitu : dimasukkannya rata-rata nilai jurusan IPA Pada tabel 7, agar terlihat bahwa nilai rata-rata SE, WA, GE dan ME dari 7 siswa lebih rendah dari 298 siswa yang masuk pada jurusan IPS.

Seperti telah dibahas pada latar belakang, bahwa penelitian terdahulu [6], [7], [8], [9] juga membahas penentuan jurusan berdasarkan nilai rata-rata raport dari siswa. Namun berbeda dengan penelitian sebelumnya, dalam penelitian ini menggunakan tes IST dalam penentuan jurusan.

Ada juga penelitian menggunakan instrument yang sama dengan penelitian ini yaitu tes IST seperti [10], namun dalam penelitian [10] juga tidak dibahas secara rinci jurusan yang dapat dibentuk dan berapa jumlah siswa pada tiap jurusan. Sehingga berbeda dengan penelitian [10], dalam penelitian ini metode terbaik dari ketiga metode digunakan untuk menentukan jumlah jurusan yang dapat terbentuk dan jumlah siswa tiap jurusan.

#### 4. KESIMPULAN

Berdasarkan data Tes Minat dan Bakat yang dilakukan pada 305 siswa di SMA N 4 Jayapura menggunakan 10 subtes IST yang dianggap sesuai dengan tujuan penelitian, maka metode yang sesuai untuk melakukan analisis terhadap data tersebut adalah metode *average linkage*. Hasil dendrogram juga memperlihatkan bahwa, jumlah jurusan yang dapat dibentuk adalah dua jurusan. Hasil analisis menunjukkan bahwa ada 7 siswa yang masuk pada jurusan IPA dan 298 siswa masuk pada jurusan IPS. Selanjutnya metode *average linkage* dapat dijadikan rujukan bagi tim tes psikologi mempercepat waktu analisis hasil tes IST. Tim tes psikologi Prodi BK UNCEN juga dapat menggunakan software Matlab R2018b untuk menganalisis data, karena model sebaran data dapat ditampilkan dalam tiga dimensi. Dengan adanya

kolaborasi bidang ilmu dan pemanfaatan software matematika yang berlisensi, diharapkan tim tes psikologi Prodi BK UNCEN dapat menjangkau seluruh SMA dan Sederajat di Provinsi Papua untuk melakukan tes psikologi dalam penentuan jurusan siswa. Dengan demikian hal ini juga berdampak positif bagi beberapa perguruan tinggi di Papua yaitu, dapat mengurangi beban kerja dosen.

#### UCAPAN TERIMA KASIH

Ucapan terima kasih kami sampaikan kepada Kementerian Pendidikan dan Kebudayaan, melalui Hibah Penelitian PNB (Penerimaan Negara Bukan Pajak) tahun anggaran 2022, pada Lembaga Penelitian dan Pengabdian kepada Masyarakat Universitas Cenderawasih.

#### DAFTAR PUSTAKA

- [1] A. R. Baqri, J. S. Putra, and K. Karimullah, "Hubungan Antara Dukungan Religius Dengan Kualitas Hidup Pada Remaja Miskin," *Indones. J. Psychol. Relig.*, 2021, doi: [10.24854/ijpr395](https://doi.org/10.24854/ijpr395).
- [2] T. P. Astuti and Y. Yulianto, "Good Governance Pengelolaan Keuangan Desa Menyongsong Berlakunya Undang-Undang No. 6 Tahun 2014," *Berk. Akunt. dan Keuang. Indones.*, 2016, doi: [10.20473/baki.v1i1.1694](https://doi.org/10.20473/baki.v1i1.1694).
- [3] P. Nastiti and Y. P. Wibisono, "Pendampingan kampanye digital melalui website dan sosial media kepada kawan kasih tumbuh (KKT)," *J. Pemberdaya. Publ. Has. ...*, 2019.
- [4] S. Winanta, Y. Oslan, G. Santoso Abstrak, and K. Kunci, "Implementasi Metode Bayesian Dalam Penjurusan Di Sma Bruderan Purworejo Studi Kasus: Sma Bruderan Purworejo," *J. Eksplor. Karya Sist. Inf. dan Sains*, 2013.
- [5] Kemendikbud, "Pedoman Peminatan Peserta Didik," *Badan Pengemb. Sumber Daya Mns. Pendidik. dan Kebud. dan Penjamin Mutu Pendidik.*, p. 74, 2013.
- [6] R. Altanova, A. Syukur, and M. A. Soeleman, "Penentuan Jurusan Siswa Sekolah Menengah Atas Disesuaikan Dengan Minat Siswa Menggunakan Algoritma Fuzzy C-Means," *J. Teknol. Inf.*, 2017.
- [7] F. N. Khasanah, "Klasifikasi Proses Penjurusan Siswa Tingkat SMA Menggunakan Data Mining," *Informatics Educ. Prof.*, 2016.
- [8] A. A. Arifiyanti, A. L. Ana, and A. D. S., "POHON KEPUTUSAN DALAM PENGKLASIFIKASIAN PENJURUSAN SISWA SEKOLAH MENENGAH ATAS ( SMA )," *Semin. Nas. Sains dan Teknol. Terap. V*, 2017.

- [9] D. Putra and A. Wibowo, "Prediksi Keputusan Minat Penjurusan Siswa SMA Yadika 5 Menggunakan Algoritma Naïve Bayes," *Pros. Semin. Nas. Ris. Dan Inf. Sci.*, 2020.
- [10] M. Mughniy, R. C. Wihandika, and B. H. Prasetyo, "Sistem Rekomendasi Psikotes untuk Penjurusan Siswa SMA menggunakan Metode Modified K-Nearest Neighbor," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, 2018.
- [11] R. Kumolohadi and miftahun ni'mah Suseno, "Intelligenz Struktur Test Dan Standard Progressive Matrices : ( Dari Konsep Inteligensi Yang Berbeda Menghasilkan Tingkat Inteligensi Yang Sama )," *J. Inov. dan Kewirausahaan*, 2012.
- [12] S. Jaya, C. Anwar, and H. Hermawan, "Sistem pemilihan program studi berdasarkan bakat, minat dan kecerdasan calon mahasiswa berbasis online," *Pros. Semin. Nas. Sains dan Teknol.*, 2017.
- [13] A. M. Sroyer, S. A. Mandowen, and F. Reba, "Analisis Cluster Penyakit Malaria Provinsi Papua Menggunakan Metode Single Linkage Dan K-Means," *J. Nas. Teknol. dan Sist. Inf.*, 2022, doi: [10.25077/teknosi.v7i3.2021.147-154](https://doi.org/10.25077/teknosi.v7i3.2021.147-154).
- [14] D. Exasanti and A. Jananto, "Analisa Hasil Pengelompokan Wilayah Kejadian Non-Kebakaran Menggunakan Agglomerative Hierarchical Clustering di Semarang," *J. Tekno Kompak*, 2021, doi: [10.33365/jtk.v15i2.1166](https://doi.org/10.33365/jtk.v15i2.1166).
- [15] R. Novidianto and A. T. R. Dani, "ANALISIS KLASTER KASUS AKTIF COVID-19 MENURUT PROVINSI DI INDONESIA BERDASARKAN DATA DERET WAKTU," *J. Apl. Stat. Komputasi Stat.*, 2020.
- [16] et al. Ghaisani, "Analisis Cluster Hirarki Untuk Pengelompokan Provinsi Di Indonesia Berdasarkan Indikator Demokrasi Indonesia Tahun 2016," *Pros. Konf. Nas. Penelit. Mat. dan Pembelajarannya IV*, 2019.
- [17] Sugiyono, "Pedoman untuk Memberikan Interpretasi Koefisien Korelasi," *Pedoman untuk memberikan Interpret. koefisien korelasi*, 2015.
- [18] A. T. R. Dani, S. Wahyuningsih, and N. A. Rizki, "Penerapan Hierarchical Clustering Metode Agglomerative pada Data Runtun Waktu," *Jambura J. Math.*, 2019, doi: [10.34312/jjom.v1i2.2354](https://doi.org/10.34312/jjom.v1i2.2354).
- [19] S. F. Mu'afa and N. Ulinuha, "Perbandingan Metode Single Linkage, Complete Linkage Dan Average Linkage dalam Pengelompokan Kecamatan Berdasarkan Variabel Jenis Ternak Kabupaten Sidoarjo," *Inf. J. Ilm. Bid. Teknol. Inf. dan Komun.*, 2019, doi: [10.25139/inform.v4i2.1696](https://doi.org/10.25139/inform.v4i2.1696).
- [20] S. I. Pratiwi, T. Widiarhari, and A. R. Hakim, "ANALISIS KLASTER METODE WARD DAN AVERAGE LINKAGE DENGAN VALIDASI DUNN INDEX DAN KOEFISIEN KORELASI COPHENETIC (Studi Kasus: Kecelakaan Lalu Lintas Berdasarkan Jenis Kendaraan Tiap Kabupaten/Kota di Jawa Tengah Tahun 2018)," *J. Gaussian*, 2019, doi: [10.14710/j.gauss.v8i4.26747](https://doi.org/10.14710/j.gauss.v8i4.26747).
- [21] B. Nugroho, "Perbandingan Aplikasi Algoritma Kernel K-Means pada Graf Bipartit dan K-Means pada Matriks Dokumen- Istilah dalam Dataset Penelitian Covid-19 RISTEKBRIN," *J. Teknol. Inf. dan Ilmu Komput.*, 2021, doi: [10.25126/jtiik.2021824365](https://doi.org/10.25126/jtiik.2021824365).
- [22] Wahyono, I. N. P. Trisna, S. L. Sariwening, M. Fajar, and D. Wijayanto, "Perbandingan Penghitungan Jarak pada K-Nearest Neighbour dalam Klasifikasi Data Tekstual," *J. Teknol. dan Sist. Komput.*, 2020.
- [23] Z. Arifin, S. Stefanus, and A. M. Soeleman, "Klasterisasi Genre Cerpen Kompas Menggunakan Agglomerative Hierarchical Clustering- Single Linkage," *J. Teknol. Inf.*, vol. 13, pp. 92–100, 2017.
- [24] A. B. E. U. Eko Pramono, "Perbandingan Metode Single Linkage dan Fuzzy C Means Untuk Pengelompokan Trafik Internet," *Respati*, 2017, doi: [10.35842/jtir.v12i2.173](https://doi.org/10.35842/jtir.v12i2.173).
- [25] R. Tjut Adek and M. Jannah, "Pencarian Kemiripan Judul Tugas Akhir Mahasiswa Dengan Menggunakan Metode Single Linkage Hierarchical," *J. SAINTEKOM*, 2018, doi: [10.33020/saintekom.v8i2.69](https://doi.org/10.33020/saintekom.v8i2.69).
- [26] N. Ulinuh and R. Veriani, "Analisis Cluster dalam Pengelompokan Provinsi di Indonesia Berdasarkan Variabel Penyakit Menular Menggunakan Metode Complete Linkage , Average Linkage dan Ward," *InfoTekJar J. Nas. Inform. dan Teknol. Jar.*, vol. 5, no. 1, pp. 101–108, 2020.
- [27] S. Ningsih, S. Wahyuningsih, and Y. N. Nasution, "Perbandingan Kinerja Metode Complete Linkage dan Average Linkage dalam Menentukan Hasil Analisis Cluster," *Pros. Semin. Sains dan Teknol. FMIPA Unmul*, 2016.
- [28] M. Paramadina, S. Sudarmin, and M. K. Aidid, "Perbandingan Analisis Cluster Metode Average Linkage dan Metode Ward (Kasus: IPM Provinsi Sulawesi Selatan)," *VARIANSI J. Stat. Its Appl. Teach. Res.*, 2019, doi: [10.35580/variansiunm9357](https://doi.org/10.35580/variansiunm9357).



- [29] E. Widodo, S. N. Mashita, and Y. G. Prasetyowati, "Perbandingan Metode Average Linkage, Complete Linkage, dan Ward'S pada Pengelompokan Kabupaten/Kota di Provinsi Jawa Tengah Berdasarkan Indikator Indeks Pembangunan Manusia," *Fakt. Exacta*, 2020, doi: [10.30998/faktorexacta.v13i2.6581](https://doi.org/10.30998/faktorexacta.v13i2.6581).
- [30] R. Wibowo, K. Nisa, A. Faisol, and E. Setiawan, "SIMULASI PEMILIHAN METODE ANALISIS CLUSTER HIRARKI AGGLOMERATIVE TERBAIK ANTARA AVERAGE LINKAGE DAN WARD PADA DATA YANG MENGANDUNG MASALAH MULTIKOLINEARITAS," *J. Siger Mat.*, 2020, doi: [10.23960/jsm.v1i2.2497](https://doi.org/10.23960/jsm.v1i2.2497).
- [31] D. Widyadhana, R. B. Hastuti, I. Kharisudin, and F. Fauzi, "Perbandingan Analisis Klaster K-Means dan Average Linkage untuk Pengklasteran Kemiskinan di Provinsi Jawa Tengah," *Prism. Pros. Semin. Nas. Mat. 4*, 2021.
- [32] Badan Pusat Statistika Provinsi Papua. "Indeks Pembangunan Manusia (IPM) Provinsi Papua Tahun 2018":  
<https://papua.bps.go.id/pressrelease/2019/05/06/424/indeks-pembangunan-manusia--ipm--provinsi-papua-tahun-2018.html> [Januari. 15, 2022].

## BIODATA PENULIS



Alvian Sroyer  
Dosen di Program Studi Matematika.  
Universitas Cenderawasih, Papua. Lulusan S2  
dari Institut Teknologi Bandung. Bidang  
penelitian tentang: Matematika Terapan.



Habel Saud  
Dosen di Program Studi Bimbingan dan  
Konseling. Universitas Cenderawasih, Papua.  
Lulusan S2 dari UNP Padang. Bidang  
penelitian tentang: Bimbingan dan Konseling.



Felix Reba  
Dosen di Program Studi Matematika.  
Universitas Cenderawasih, Papua. Lulusan S2  
dari Universitas Gadjja Mada. Bidang penelitian  
tentang: Matematika Terapan, Statistika  
Komputasi dan Terapan .