ANALISIS *K-MEANS CLUSTER* INDEKS PEMBANGUNAN MANUSIA DI INDONESIA

Tugas Akhir

disusun sebagai salah satu syarat untuk memperoleh gelar Ahli Madya

Program Studi Statistika Terapan dan Komputasi

oleh

Muchlas Santoso

4112315018

JURUSAN MATEMATIKA

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

UNIVERSITAS NEGERI SEMARANG

2019

PERNYATAAN

Saya menyatakan bahwa Tugas Akhir ini bebas plagiat dan belum pernah diajukan untuk memperoleh gelar diploma di suatu instansi perguruan tinggi, dan sepanjang pengetahuan saya, tidak terdapat karya atau pendapat yang pernah diterbitkan oleh orang lain, kecuali yang secara tertulis dirujuk dalam Tugas Akhir ini dan disebutkan dalam daftar pustaka.

Semarang, 30 Agustus 2019

Muchlas Santoso

4112315018

PENGESAHAN

Tugas Akhir yang berjudul

Analisis *K-Means Cluster* Indeks Pembangunan Manusia di Indonesia

disusun oleh

Nama : Muchlas Santoso

NIM : 4112315018

telah dipertahankan di hadapan sidang Panitia Ujian Tugas Akhir FMIPA Universitas Negeri Semarang pada tanggal 8 Oktober 2019

Panitia

Ketua

Dr. Sugianto, M.Si.
NIP 196102191993031001

Sekretaris

Dr. Mulyono, M.Si.
NIP 197009021997021001

Pembimbing / Penguji II

Dr. Walid, S.Pd., M.Si.
NIP 197408192001121001

Penguji I

Drs. Sugiman, M.Si.
NIP 196401111989011001

MOTTO DAN PERSEMBAHAN

Motto

Dengan Ilmu, hidup akan menjadi lebih Mudah

Dengan Cinta, hidup akan menjadi lebih Indah

Dengan Iman, hidup akan menjadi lebih Terarah

Persembahan

Tugas Akhir ini saya persembahkan untuk:

- Bapak saya, Sulasno dan Ibu saya, Jamiatul Islamiyah, terima kasih untuk kasih sayang, pengorbanan, dukungan serta do'anya.
- Adik saya, Mutiah Syifa Isnaini.
- Seseorang yang tertulis di Lawhul Mahfudz.
- Seseorang yang memaksa saya untuk segera menyelesaikan TA.
- Semua orang yang telah memberikan bantuan, motivasi, dan semangat, serta orang yang hanya memberikan sindiran.
- Teman-teman Dolan Bareng yang selalu hadir dan membantu dalam penyusunan TA.
- Teman-teman Mathematics Computing Club (MCC) dengan segala perbendaharaan katanya.
- Teman-teman Staterkom angkatan 2015 yang saling menyemangati.
- Mahasiswa Jurusan Matematika Universitas Negeri Semarang.
- Universitas Negeri Semarang.

KATA PENGANTAR

Puji syukur kehadirat Allah SWT karena atas rahmat, taufik, hidayah dan inayah-Nya saya dapat menyelesaikan Tugas Akhir dengan judul "**Analisis *K-Means Cluster* Indeks Pembangunan Manusia di Indonesia**". Sholawat serta salam Allah SWT senantiasa tercurahkan kepada junjungan kita Nabi Muhammad SAW, yang kita nantikan safa'atnya di hari kiamat nanti, Aamiin.

Tugas Akhir ini disusun sebagai salah satu syarat memperoleh gelar Ahli Madya Program Studi Statistika Terapan dan Komputasi.

Penulis menyadari dalam penulisan Tugas Akhir ini tidak lepas dari bantuan, motivasi, serta dorongan semangat dari beberapa pihak. Oleh karena itu penulis menyampaikan rasa hormat dan terima kasih kepada:

1. Bapak saya, Sulasno, dan Ibu saya, Jamiatul Islamiyah yang telah memberikan semangat, pengorbanan, kasih sayang, dan do'a selama penyusunan Tugas Akhir sampai Tugas Akhir ini selesai disusun.
2. Bapak Prof. Dr. Fathur Rokhman, M.Hum., Rektor Universitas Negeri Semarang,
3. Bapak Dr. Sugianto, M.Si., Dekan Fakultas MIPA Universitas Negeri Semarang,
4. Bapak Dr. Mulyono, M.Si., Ketua Jurusan Matematika Fakultas MIPA Universitas Negeri Semarang,
5. Bapak Dr. Iqbal Kharisudin, S.Pd., M.Sc., Koordinator Program Studi Statistika Terapan dan Komputasi Jurusan Matematika Fakultas MIPA Universitas Negeri Semarang,
6. Bapak Dr. Walid, S.Pd., M.Si., Dosen Pembimbing Tugas Akhir yang telah memberikan bimbingan dan masukannya selama penyusunan Tugas Akhir,
7. Bapak Drs. Sugiman, M.Si. Dosen Penguji Tugas Akhir yang telah memberikan kritik dan saran yang membangun untuk perbaikan Tugas Akhir saya, serta menambah pengetahuan saya dalam materi Tugas Akhir yang saya tulis.

8. Teman-teman yang telah memberikan motivasi, semangat, dan do’a selama mengerjakan Tugas Akhir, dan
9. Semua pihak yang telah membantu dalam proses penyusunan Tugas Akhir.

Penulis berharap agar Tugas Akhir ini dapat berguna dan bermanfaat bagi penulis pada khususnya dan pembaca pada umumnya.

Penulis menyadari sepenuhnya bahwa dalam penyusunan Tugas Akhir ini masih terdapat kekurangan dan jauh dari kata sempurna, untuk itu kritik dan saran yang bersifat konstruktif sangat saya harapkan demi perbaikan dan kesempurnaan Tugas Akhir ini.

Semarang, 30 Agustus 2019

Penulis

ABSTRAK

Santoso, Muchlas. 2019. *Analisis K-Means Cluster Indeks Pembangunan Manusia di Indonesia*. Tugas Akhir, Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Negeri Semarang. Pembimbing Utama Dr. Walid, S.Pd., M.Si.

Pembangunan manusia dalam suatu negara diperlukan untuk menciptakan kesejahteraan dan membuat perkembangan dalam suatu negara sehingga perlu dilihat terlebih dahulu kualitas sumber daya manusia dalam negara tersebut. Ukuran yang dijadikan standar keberhasilan pembangunan manusia dalam sebuah negara adalah Indeks Pembangunan Manusia (IPM) atau *Human Development Index* (HDI). Konsep pembangunan manusia diukur dengan menggunakan pendekatan tiga dimensi dasar manusia, yaitu umur panjang dan sehat, pengetahuan, dan standar hidup yang layak. Dimensi umur panjang dan sehat diwakili oleh indikator harapan hidup saat lahir. Dimensi pengetahuan diwakili oleh indikator harapan lama sekolah dan rata-rata lama sekolah. Sementara itu, dimensi standar hidup layak diwakili oleh pengeluaran per kapita. Tujuan dari penelitian ini adalah untuk menjelaskan karakteristik masing-masing *cluster* yang diperoleh dari hasil pengelompokkan provinsi di Indonesia berdasarkan indikator penyusun IPM tahun 2017 menggunakan metode *K-Means Cluster*, menjelaskan model diskriminan 2 faktor yang terbentuk, serta untuk mengetahui ketepatan hasil dalam mengklasifikasi kasus dari model diskriminan yang terbentuk.

Metode yang digunakan dalam penelitian ini adalah Analisis *K-Means Cluster* dan Analisis Diskriminan. Analisis *K-Means Cluster* digunakan untuk mengelompokkan 34 provinsi di Indonesia berdasarkan indikator penyusun IPM tahun 2017. Hasil pengelompokkan dari Analisis *K-Means Cluster*, kemudian dianalisis lebih lanjut menggunakan Analisis Diskriminan untuk menguji keakuratan hasil dari pengelompokkan menggunakan Analisis *K-Means Cluster*. Analisis Diskriminan juga digunakan untuk mengetahui faktor yang paling berpengaruh dalam penyusunan nilai Indeks Pembangunan Manusia.

Hasil analisis dari penelitian ini adalah terbentuknya 2 kelompok (*cluster*) hasil dari Analisis *K-Means Cluster*. Karakteristik *Cluster* 1 berisi provinsi dengan umur harapan hidup (UHH) rendah, harapan lama sekolah (HLS) rendah, rata-rata lama sekolah (RLS) rendah, dan pengeluaran disesuaikan yang rendah. Karakteristik *Cluster* 2 berisi provinsi dengan umur harapan hidup (UHH) tinggi, harapan lama sekolah (HLS) tinggi, rata-rata lama sekolah (RLS) tinggi, dan pengeluaran disesuaikan yang tinggi. Model diskriminan 2 faktor yang adalah $Y = -11{,}642 + 1{,}409\ RLS$. Jadi, indikator yang paling berpengaruh dalam penyusunan nilai IPM adalah rata-rata lama sekolah. Ketepatan hasil dalam mengklasifikasi kasus dari model diskriminan yang terbentuk adalah sebesar 91.20%. Jadi, dapat disimpulkan bahwa hasil pengelompokkan menggunakan metode K-Means Cluster mempunyai tingkat keakuratan yang tinggi.

Kata kunci: Analisis Diskriminan, Analisis *K-Means Cluster*, Umur Panjang dan Sehat, Pengetahuan, Standar Hidup Layak, Indeks Pembangunan Manusia.

DAFTAR ISI

HALAMAN JUDUL i
PERNYATAAN ii
PENGESAHAN iii
MOTTO DAN PERSEMBAHAN iv
KATA PENGANTAR v
ABSTRAK vii
DAFTAR ISI viii
DAFTAR TABEL x
DAFTAR GAMBAR xi
DAFTAR LAMPIRAN xii
BAB I PENDAHULUAN 1
1.1 Latar Belakang 1
1.2 Rumusan Masalah 4
1.3 Batasan Masalah 4
1.4 Tujuan Penelitian 4
1.5 Manfaat Penelitian 4
1.6 Sistematika Penulisan 5
BAB II LANDASAN TEORI 7
2.1 Indeks Pembangunan Manusia 7
2.1.1 Pengertian Indeks Pembangunan Manusia 7
2.1.2 Manfaat Indeks Pembangunan Manusia 9
2.1.3 Dimensi Dasar Indeks Pembangunan Manusia 9
2.2 Analisis *Cluster* 17
2.2.1 Definisi Analisis *Cluster* 17
2.2.2 Hal-Hal Pokok Tentang Analisis *Cluster* 17
2.2.3 Proses Dasar dari Analisis *Cluster* 18
2.2.4 Metode Analisis *Cluster* 20
2.2.5 *K-Means Cluster* 24
2.3 Analisis Diskriminan 26
2.3.1 Definisi Analisis Diskriminan 26
2.3.2 Kegunaan Analisis Diskriminan 26

2.3.3 Tujuan Analisis Diskriminan 27
2.3.4 Jumlah Sampel Diskriminan 27
2.3.5 Asumsi pada Analisis Diskriminan 28
2.3.6 Proses Dasar Analisis Diskriminan 28
BAB III METODOLOGI PENELITIAN 31
3.1 Jenis dan Sumber Data 31
3.2 Variabel Penelitian 31
3.3 Teknik Pengumpulan Data 31
3.4 Teknik Analisis Data 32
3.4.1 Tahapan Analisis *K-Means Cluster* 32
3.4.2 Tahapan Analisis Diskriminan 33
BAB IV HASIL DAN PEMBAHASAN 34
4.1 Hasil Penelitian 34
4.1.1 Analisis Deskriptif 34
4.1.2 Uji Asumsi Analisis *K-Means Cluster* 35
4.1.3 Proses Awal *Clustering* 37
4.1.4 Proses Iterasi 38
4.1.5 Hasil akhir proses *Cluster*ing 38
4.1.6 Melihat perbedaan variabel pada *Cluster* yang terbentuk 40
4.1.7 Jumlah anggota di setiap *Cluster* 41
4.1.8 Penafsiran komposisi pada *Cluster* 41
4.1.9 Uji Asumsi Analisis Diskriminan 44
4.1.10 Fungsi Diskriminan Canonical 50
4.1.11 Fungsi Diskriminan Fisher 50
4.1.12 Ketepatan Hasil Klasifikasi 51
4.2 Pembahasan 51
BAB V PENUTUP 57
5.1 Simpulan 57
5.2 Saran 58
DAFTAR PUSTAKA 59
LAMPIRAN 63

DAFTAR TABEL

Tabel 2. 1 Pertumbuhan Indeks Pembangunan Manusia, 2010-2017 9
Tabel 2. 2 Pertumbuhan Umur Harapan Hidup di Indonesia, 2010-2017 10
Tabel 2. 3 Umur Harapan Hidup Menurut Provinsi di Indonesia Tahun 2017..... 11
Tabel 2. 4 Pertumbuhan Harapan Lama Sekolah (HLS) dan Rata-Rata Lama Sekolah (RLS) di Indonesia, 2010-2017 13
Tabel 2. 5 Harapan Lama Sekolah (HLS) dan Rata-Rata Lama Sekolah (RLS) Menurut Provinsi di Indonesia Tahun 2017 14
Tabel 2. 6 Pertumbuhan Harapan Lama Sekolah (HLS) dan Rata-Rata Lama Sekolah (RLS) di Indonesia, 2010-2017 15
Tabel 2. 7 Pengeluaran per Kapita yang Disesuaikan Menurut Provinsi di Indonesia Tahun 2017 16
Tabel 4. 1 *Descriptive Statistics* 34
Tabel 4. 2 *One-Sample Kolmogorov-Smirnov Test* 35
Tabel 4. 3 *Correlations* 36
Tabel 4. 4 KMO *and Bartlett's Test* 37
Tabel 4. 5 *Initial Cluster Centers* 37
Tabel 4. 6 *Iteration History*[a] 38
Tabel 4. 7 *Final Cluster Centers* 38
Tabel 4. 8 *ANOVA* 40
Tabel 4. 9 *Number of Cases in each Cluster* 41
Tabel 4. 10 Pembagian *Cluster* Menurut Provinsi di Indonesia 41
Tabel 4. 11 *One-Sample Kolmogorov-Smirnov Test* 45
Tabel 4. 12 *Coefficients*[a] 45
Tabel 4. 13 *Tests of Equality of Group Means* 46
Tabel 4. 14 *Test Results* 47
Tabel 4. 15 *Canonical Discriminant Function Coefficients* 50
Tabel 4. 16 *Classification Function Coefficients* 50
Tabel 4. 17 *Classification Results*[a,c] 51

DAFTAR GAMBAR

Gambar 2. 1 Indeks Pembangunan Manusia beserta Indikatornya 8
Gambar 4. 1 *Box Plot* Umur Harapan Hidup 48
Gambar 4. 2 *Box Plot* Harapan Lama Sekolah 48
Gambar 4. 3 *Box Plot* Rata-Rata Lama Sekolah 49
Gambar 4. 4 *Box Plot* Pengeluaran Disesuaikan..... 49

DAFTAR LAMPIRAN

Lampiran 1. Indikator Penyusun Indeks Pembangunan Manusia Tahun 2017..... 64
Lampiran 2. Analisis Deskriptif..... 68
Lampiran 3. Uji Asumsi Analisis *K-Means Cluster* 69
Lampiran 4. Analisis *K-Means Cluster*..... 70
Lampiran 5. Uji Asumsi Analisis Diskriminan..... 73
Lampiran 6. Analisis Diskriminan 76

BAB I
PENDAHULUAN

1.1 Latar Belakang

Pembangunan manusia dalam suatu negara diperlukan untuk menciptakan kesejahteraan dan membuat perkembangan dalam suatu negara sehingga perlu dilihat terlebih dahulu kualitas sumber daya manusia dalam negara tersebut. Diharapkan dengan meningkatnya kualitas sumber daya manusia dapat meningkatkan kualitas hidup manusia pada negara tersebut sehingga membuat negara tersebut tidak terjadi kesenjangan sosial.

Martabat manusia seharusnya menjadi tujuan hakiki pembangunan sebuah negara. Demikian, rekomendasi lembaga dunia *United Nations Development Programme* (UNDP) yang juga menyatakan bahwa manusia yang bermartabat adalah manusia yang dapat menikmati umur panjang, sehat, dan menjalankan kehidupan yang produktif (*Human Development Report* 1990). Karenanya, hingga kini ukuran yang dijadikan standar keberhasilan pembangunan manusia sebuah negara adalah Index Pembangunan Manusia (IPM) atau *Human Development Index* (HDI). Konsep pembangunan manusia itu sendiri pada dasarnya memiliki makna yang sangat luas. Konsep ini mencakup semua dimensi dasar yang dimiliki oleh manusia. Namun, ide dasar dari konsep pembangunan manusia pada intinya cukup sederhana, yaitu menciptakan pertumbuhan positif dalam bidang ekonomi, sosial, politik, budaya, dan lingkungan, serta perubahan dalam kesejahteraan manusia. Oleh karena itu, manusia harus diposisikan sebagai kekayaan bangsa yang sesungguhnya. (Badan Pusat Statistik, 2018)

Pembangunan ekonomi adalah suatu proses kenaikan pendapatan total dan pendapatan perkapita dengan memperhitungkan adanya pertambahan penduduk dan disertai dengan perubahan fundamental dalam struktur ekonomi suatu daerah dan pemerataan pendapatan bagi suatu penduduk suatu daerah. Cara paling mudah untuk mengartikan pembangunan ekonomi adalah dimana pertumbuhan ekonomi ditambah dengan perubahan. Artinya, ada tidaknya pembangunan ekonomi suatu daerah pada suatu tahun tertentu tidak saja diukur dari kenaikan produksi barang

dan jasa yang berlaku dari tahun ke tahun, tetapi juga perlu diukur dari perubahan lain yang berlaku dari berbagai aspek kegiatan ekonomi seperti perkembangan pendidikan, perkembangan teknologi, peningkatan dalam kesehatan, peningkatan dalam infrastruktur yang tersedia dan peningkatan dalam pendapatan dan kemakmuran masyarakat (Sukirno, 2006).

Menurut (Badan Pusat Statistik, 2018) indikator penting untuk mengukur keberhasilan dalam upaya membangun kualitas hidup manusia adalah indeks pembangunan manusia. (Pratowo, 2013) juga menjelaskan bahwa IPM merupakan indeks komposit yang dihitung sebagai rata-rata sederhana dari tiga indeks dari dimensi yang menggambarkan kemampuan dasar manusia dalam memperluas pilihan-pilihan. Dimana 3 indeks tersebut adalah indeks harapan hidup, pendidikan, dan standar hidup layak.

Berdasarkan hasil penghitungan, BPS mencatat IPM Indonesia pada tahun 2017 sebesar 70,81. Angka IPM dihitung dari agregasi tiga dimensi, yaitu umurpanjang dan hidup sehat, pengetahuan, serta standar hidup layak. Setiap dimensi diwakili oleh indikator. Dimensi umur panjang dan hidup sehat diwakili oleh indikator umur harapan hidup saat lahir. Sementara itu, rata-rata lama sekolah dan harapan lama sekolah merupakan indikator yang mewakili dimensi pengetahuan. Terakhir, dimensi standar hidup layak Indonesia diwakili oleh indikator pengeluaran per kapita yang disesuaikan. Berdasarkan data pada tahun 2017, umur harapan hidup saat lahir di Indonesia sudah mencapai 71,06 tahun. Angka ini menunjukkan bahwa hidup bayi yang baru lahir dapat bertahan hidup hingga usia 71,06 tahun. Dari sisi pendidikan, penduduk Indonesia yang berusia 25 tahun ke atas rata-rata sudah menempuh 8,10 tahun masa sekolah atau telah menyelesaikan kelas VIII. Selain itu, rata-rata anak usia 7 tahun yang mulai bersekolah, diperkirakan dapat mengenyam pendidikan hingga 12,85 tahun atau setara dengan kelas XII atau tamat SMA. Pada aspek ekonomi, pengeluaran per kapita sudah mencapai Rp10.664.000 per kapita per tahun.

BPS juga mencatat sejak tahun 2010, pembangunan manusia di Indonesia memperlihatkan perkembangan yang positif. Setiap tahun, IPM Indonesia meningkat rata-rata 0,89 persen per tahun. Pada kurun waktu tujuh tahun, telah

terjadi kenaikan IPM hingga 4,28 poin. Perkembangan ini menunjukkan semakin membaiknya pembangunan manusia secara umum di Indonesia. Pada tahun 2017, IPM di Indonesia sudah mencapai 70,81, meningkat 0,63 poin dibanding tahun sebelumnya. Hal ini berarti IPM di Indonesia tumbuh 0,90 persen pada periode 2016-2017.

Berdasarkan penjabaran dari latar belakang di atas, penulis ingin melakukan pengelompokkan provinsi di Indonesia menggunakan analisis *K-Means Cluster* berdasarkan indikator penyusun Indeks Pembangunan Manusia tahun 2017, kemudian akan dianalisis lebih lanjut menggunakan analisis diskriminan untuk menguji keakuratan hasil pengelompokan *K-Means Cluster*. Selanjutnya, penulis menuliskannya dalam bentuk Tugas Akhir dengan judul "**Analisis *K-Means Cluster* Indeks Pembangunan Manusia di Indonesia**".

1.2 Rumusan Masalah

Berdasarkan latar belakang tersebut, maka dapat dirumuskan masalah sebagai berikut.

1. Berapa banyak *cluster* (kelompok) yang terbentuk dan bagaimana karakteristiknya?
2. Bagaimana model diskriminan yang terbentuk dari Analisis *K-Means Cluster*?
3. Bagaimana ketepatan hasil dalam mengklasifikasi kasus dari model diskriminan yang terbentuk?

1.3 Batasan Masalah

Untuk menghindari pembahasan yang terlalu melebar, terdapat beberapa batasan pada penelitian ini, antara lain:

1. Data yang digunakan adalah data indeks pembangunan manusia tahun 2017.
2. Untuk sampel data yang dianalisis adalah provinsi di Indonesia.

1.4 Tujuan Penelitian

Berdasarkan latar belakang dan rumusan masalah di atas maka tujuan yang ingin dicapai dalam penelitian in ini adalah:

a. Mengetahui banyaknya *cluster* yang terbentuk beserta karakteristiknya.
b. Mengetahui model diskriminan yang terbentuk dari hasil pengelompokkan Analisis *K-Means Cluster*.
c. Mengetahui ketepatan hasil dalam mengklasifikasi kasus dari model diskriminan yang terbentuk.

1.5 Manfaat Penelitian

Adapun manfaat penelitian ini adalah sebagai berikut.

1. Bagi Mahasiswa
   a. Menambah wawasan mengenai Indeks Pembangunan Manusia di Provinsi-provinsi yang ada di Indonesia.
   b. Menerapkan ilmu yang telah diperoleh selama perkuliahan

c. Memberikan pengetahuan kepada mahasiswa bahwa ilmu statistik dapat diterapkan di bidang pemerintahan.

2. Bagi Jurusan Matematika

a. Sebagai bahan referensi bagi pihak perpustakaan dan bahan bacaan yang dapat menambah ilmu pengetahuan bagi para pembaca.

b. Dapat dijadikan sebagai bahan studi kasus bagi pembaca dan acuan bagi mahasiswa yang akan melakukan analisis yang berkaitan tentang Indeks Pembangunan Manusia.

3. Bagi Pemerintahan

Sebagai bahan studi kasus ataupun acuan untuk menangani masalah yang ada di dalam provinsi-provinsi di Indonesia mengenai Indeks Pembangunan Manusia dan yang mempengaruhinya.

1.6 Sistematika Penulisan

Sistematika penulisan dari Tugas Akhir ini terdiri menjadi 3 bagian, yaitu bagian awal (pendahuluan), bagian isi, dan bagian akhir (penutup) yang masing-masing akan diuraikan sebagai berikut.

Bagian awal berisi halaman judul, pernyataan, pengesahan, motto dan persembahan, kata pengantar, abstrak, daftar isi, daftar tabel, daftar gambar dan daftar lampiran.

Bagian isi berisi laporan penelitian yang terdiri atas pendahuluan, landasan teori, metodologi penelitian, hasil dan pembahasan, dan penutup yang disusun menjadi 5 bab dengan rincian sebagai berikut.

BAB I PENDAHULUAN

Bab ini berisi latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian dan sistematika penulisan.

BAB II LANDASAN TEORI

Bab ini berisi teori – teori yang akan digunakan sebagai data pendukung dalam penyelesaian masalah dalam penelitian ini.

BAB III METODOLOGI PENELITIAN

Bab ini berisi tentang metode / analisis yang digunakan dalam Tugas Akhir ini, yaitu analisis *k-means cluster* dilanjut dengan analisis diskriminan beserta tahapan-tahapan dari masing-masing analisis tersebut.

BAB IV HASIL DAN PEMBAHASAN

Bab ini berisi hasil dari analisis yang telah dilakukan lengkap beserta interpretasi dan pembahasannya.

BAB IV PENUTUP

Bab ini berisi simpulan dan saran. Simpulan merupakan hasil pembahasan dari bab-bab sebelumnya dan merupakan jawaban dari rumusan masalah pada Tugas Akhir ini.

Bagian akhir berisi daftar pustaka yang digunakan sebagai acuan yang digunakan sebagai acuan dan lampiran – lampiran untuk melengkapi uraian pada bagian isi.

BAB II
LANDASAN TEORI

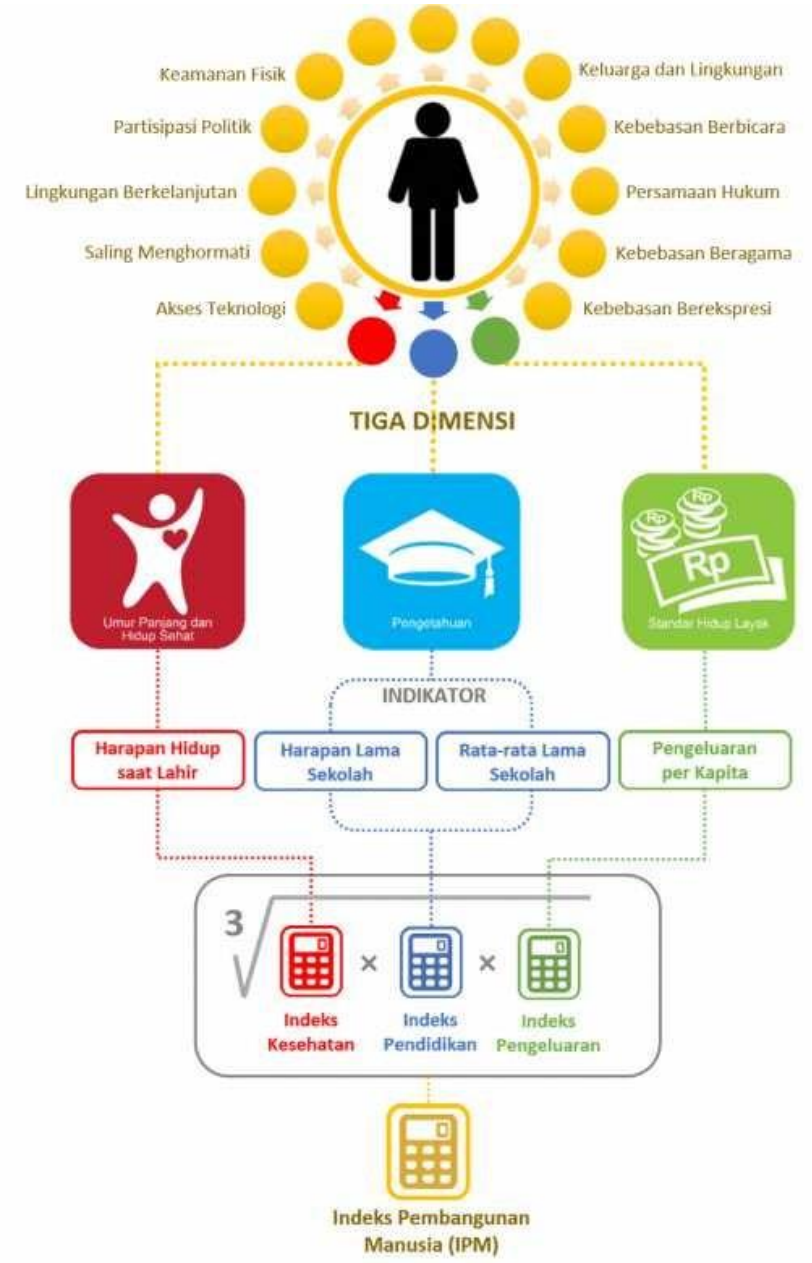
2.1 Indeks Pembangunan Manusia

2.1.1 Pengertian Indeks Pembangunan Manusia

Menurut Badan Pusat Statistik (2018), Indeks Pembangunan Manusia menjelaskan bagaimana penduduk dapat mengakses hasil pembangunan dalam memperoleh pendapatan, kesehatan, pendidikan, dan sebagainya.

Menurut BPS Kabupaten Klaten (2017), IPM adalah suatu indeks komposit yang menunjukkan capaian keberhasilan pembangunan manusia di suatu wilayah. IPM menjelaskan bagaimana penduduk dapat mengakses hasil pembangunan dalam memperoleh pendapatan, kesehatan, pendidikan dan sebagainya. IPM diperkenalkan oleh UNDP pada tahun 1990 serta dipublikasikan secara berkala dalam laporan tahunan *Human Development Report* (HDR).

Menurut Badan Pusat Statistik (2018), konsep pembangunan manusia itu sendiri pada dasarnya memiliki makna yang sangat luas. Konsep ini mencakup semua dimensi dasar yang dimiliki oleh manusia. Namun, ide dasar dari konsep pembangunan manusia pada intinya cukup sederhana, yaitu menciptakan pertumbuhan positif dalam bidang ekonomi, sosial, politik, budaya, dan lingkungan, serta perubahan dalam kesejahteraan manusia. Oleh karena itu, manusia harus diposisikan sebagai kekayaan bangsa yang sesungguhnya. Konsep pembangunan manusia diukur dengan menggunakan pendekatan tiga dimensi dasar manusia, yaitu umur panjang dan hidup sehat, pengetahuan, dan standar hidup yang layak. Dimensi umur panjang dan sehat diwakili oleh indikator umur harapan hidup saat lahir. Dimensi pengetahuan diwakili oleh indikator harapan lama sekolah dan rata-rata lama sekolah. Sementara itu, dimensi standar hidup layak diwakili oleh pengeluaran per kapita yang disesuaikan. Ketiga dimensi ini terangkum dalam suatu indeks komposit yang membentuk IPM.



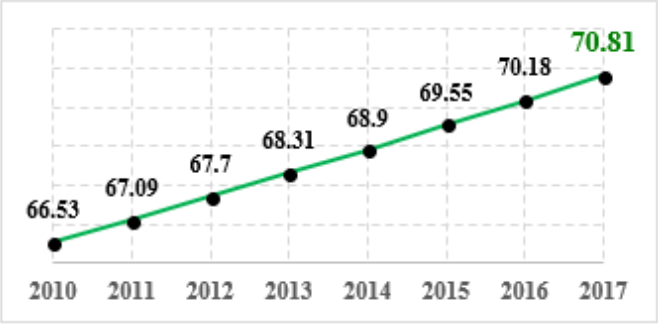
Gambar 2. 1 Indeks Pembangunan Manusia beserta Indikatornya

Sejak tahun 2010, pembangunan manusia di Indonesia memperlihatkan perkembangan yang positif. Setiap tahun, IPM Indonesia meningkat rata-rata 0,89 persen per tahun. Dalam kurun waktu tujuh tahun, telah terjadi kenaikan IPM hingga 4,28 poin. Perkembangan ini menunjukkan semakin membaiknya pembangunan manusia secara umum di Indonesia. Pada tahun 2017, IPM di Indonesia sudah mencapai 70,81, meningkat 0,63 poin dibanding tahun

sebelumnya. Hal ini berarti IPM di Indonesia tumbuh 0,90 persen pada periode 2016-2017 (BPS, 2017).

Tabel 2.1 Pertumbuhan Indeks Pembangunan Manusia, 2010-2017

Tahun	Pertumbuhan
2010-2011	0,84
2011-2012	0,90
2012-2013	0,86
2013-2014	0,87
2014-2015	0,88
2015-2016	0,89
2016-2017	0,90



Sumber: Badan Pusat Statistik

2.1.2 Manfaat Indeks Pembangunan Manusia

Badan Pusat Statistik (2018), juga menjelaskan bahwa IPM juga memiliki beberapa manfaat, diantaranya adalah:

a. IPM merupakan indikator penting untuk mengukur keberhasilan dalam upaya membangun kualitas hidup manusia (masyarakat/penduduk).

b. IPM dapat menentukan peringkat atau level pembangunan suatu wilayah/negara.

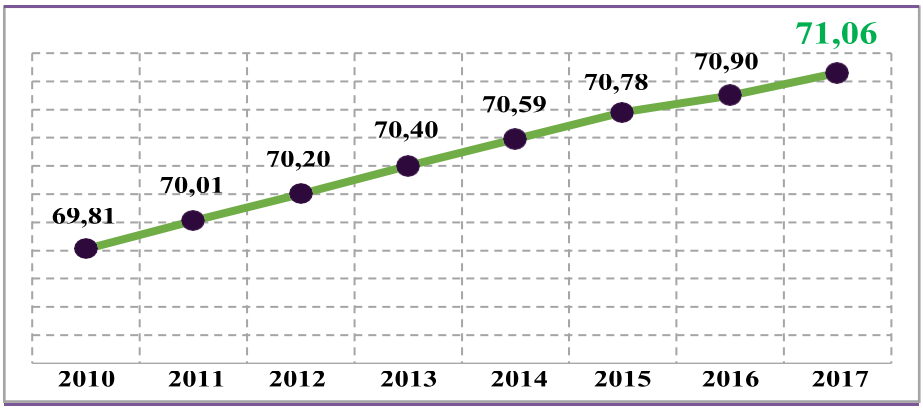
c. Bagi Indonesia, IPM merupakan data strategis karena selain sebagai ukuran kinerja Pemerintah, IPM juga digunakan sebagai salah satu alokator penentuan Dana Alokasi Umum (DAU).

2.1.3 Dimensi Dasar Indeks Pembangunan Manusia

2.1.3.1 Umur Panjang dan Hidup Sehat

Indikator umur (angka) harapan hidup saat lahir merupakan indikator yang mewakili dimensi umur panjang dan hidup sehat dalam penghitungan IPM. Umur harapan hidup saat lahir merupakan indikator yang mencerminkan derajat kesehatan masyarakat pada suatu wilayah, baik dari sarana prasarana, akses, maupun kualitas kesehatan. Secara tidak langsung, peningkatan umur harapan hidup menunjukkan derajat kesehatan masyarakat semakin baik dalam semua aspek kesehatan (Badan Pusat Statistik, 2018).

Tabel 2.2 Pertumbuhan Umur Harapan Hidup di Indonesia, 2010-2017



Sumber: Badan Pusat Statistik

Angka harapan hidup saat lahir adalah rata-rata perkiraan banyak tahun yang dapat ditempuh oleh seseorang sejak lahir. AHH mencerminkan derajat kesehatan suatu masyarakat. AHH dihitung dari hasil sensus dan survei kependudukan. (BPS Kabupaten Klaten, 2017).

Umur harapan hidup saat lahir Indonesia terus menunjukkan peningkatan dari tahun 2010 hingga 2017. Hal ini menunjukkan harapan bayi yang baru lahir untuk hidup semakin besar. Secara tidak langsung, hal ini merupakan cerminan semakin baiknya derajat kesehatan masyarakat di Indonesia. Banyak hal dapat mendorong peningkatan derajat kesehatan masyarakat. Kemajuan teknologi di bidang kesehatan, peningkatan sarana dan prasarana kesehatan, serta kepedulian masyarakat terhadap gaya hidup sehat yang meningkat turut berperan dalam memperbaiki kualitas kesehatan masyarakat (Badan Pusat Statistik, 2018).

Tabel 2. 3 Umur Harapan Hidup (UHH) Menurut Provinsi di Indonesia Tahun 2017

Provinsi	UHH (tahun)
Aceh	69,52
Sumatera Utara	68,37
Sumatera Barat	68,78
Riau	70,99
Jambi	70,76
Sumatera Selatan	69,18
Bengkulu	68,59
Lampung	69,95
Kep. Bangka Belitung	69,95
Kepulauan Riau	69,48
DKI Jakarta	72,55
Jawa Barat	72,47
Jawa Tengah	74,08
D I Yogyakarta	74,74
Jawa Timur	70,80
Banten	69,49
Bali	71,46
Nusa Tenggara Barat	65,55
Nusa Tenggara Timur	66,07
Kalimantan Barat	69,92
Kalimantan Tengah	69,59
Kalimantan Selatan	68,02
Kalimantan Timur	73,70
Kalimantan Utara	72,47
Sulawesi Utara	71,04
Sulawesi Tengah	67,32
Sulawesi Selatan	69,84
Sulawesi Tenggara	70,47
Gorontalo	67,14
Sulawesi Barat	64,34
Maluku	65,40
Maluku Utara	67,54
Papua Barat	65,32
Papua	65,14
INDONESIA	**71,06**

2.1.3.2 Pengetahuan

Dimensi pengetahuan dibentuk dari dua indikator, yaitu harapan lama sekolah dengan rata-rata lama sekolah. Cakupan dalam menghitung harapan lama sekolah adalah pendidikan penduduk dari usia 7 tahun ke atas, sementara cakupan penduduk untuk menghitung rata-rata lama sekolah yaitu dari usia 25 tahun ke atas. Kedua indikator ini diagregasikan menjadi indeks pendidikan dalam penghitungan IPM.

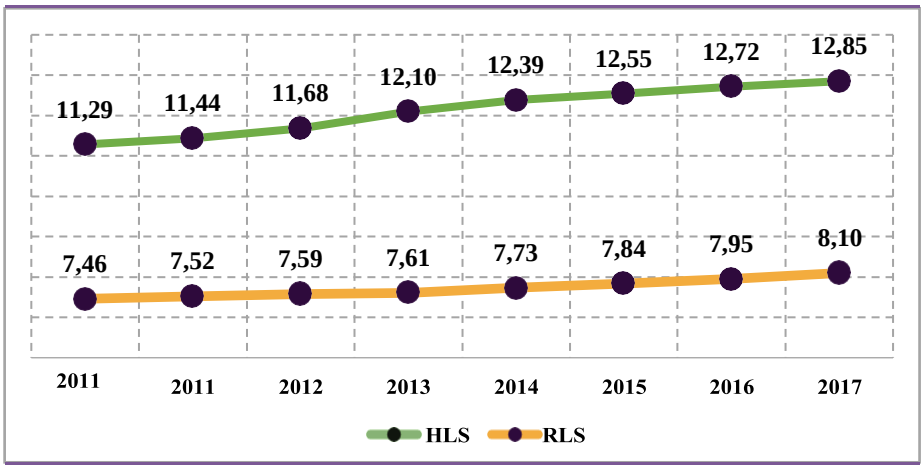
Menurut BPS Kabupaten Klaten (2017), angka harapan lama sekolah adalah lamanya sekolah (dalam tahun) yang diharapkan akan dirasakan oleh anak umur tertentu di masa mendatang. Diasumsikan bahwa peluang anak tersebut akan tetap bersekolah pada umur-umur berikutnya sama dengan peluang penduduk yang bersekolah per jumlah penduduk untuk umur yang sama saat ini. Angka Harapan Lama Sekolah dihitung untuk penduduk berusia 7 tahun ke atas. HLS dapat digunakan untuk mengetahui kondisi pembangunan sistem pendidikan di berbagai jenjang yang ditunjukkan dalam bentuk lamanya pendidikan (dalam tahun) yang diharapkan dapat dicapai oleh setiap anak.

Rata-rata lama sekolah adalah jumlah tahun yang digunakan oleh penduduk dalam menjalani pendidikan formal. Diasumsikan bahwa dalam kondisi normal rata-rata lama sekolah suatu wilayah tidak akan turun. Cakupan penduduk yang dihitung dalam penghitungan rata-rata lama sekolah adalah penduduk berusia 25 tahun ke atas (BPS Kabupaten Klaten, 2017).

Menurut BPS (2018), selama Periode 2010 hingga 2017, pengetahuan masyarakat Indonesia semakin meningkat. Hal ini sejalan dengan indikator harapan lama sekolah dan rata-rata lama sekolah yang menunjukkan peningkatan. Meskipun demikian, rata-rata lama sekolah cenderung tumbuh lebih lambat dibandingkan dengan harapan lama sekolah. Hal ini wajar karena rata-rata lama sekolah menggambarkan indikator *output* pembangunan jangka panjang, sedangkan harapan lama sekolah menggambarkan partisipasi sekolah penduduk umur 7 tahun ke atas. Perlu diketahui, indikator ini merupakan indikator proses pembangunan sebagai ukuran keberhasilan program-program pendidikan jangka pendek. Lebih jauh, rata-rata lama sekolah dan harapan lama sekolah dapat memberikan gambaran

tentang capaian (*stock*) dan penambahan (*flow*) sumber daya manusia berkualitas di suatu wilayah.

Tabel 2. 4 Pertumbuhan Harapan Lama Sekolah (HLS) dan Rata-Rata Lama Sekolah (RLS) di Indonesia, 2010-2017



Sumber: Badan Pusat Statistik

Harapan lama sekolah dan rata-rata lama sekolah di Indonesia menunjukkan tren yang meningkat selama periode 2010 hingga 2017. Secara rata-rata, harapan lama sekolah usia 7 tahun tumbuh sebesar 1,87 persen per tahun selama tahun 2010-2017. Sementara itu, rata-rata lama sekolah hanya tumbuh sebesar 1,18 persen per tahun dalam kurun waktu yang sama (Badan Pusat Statistik, 2018).

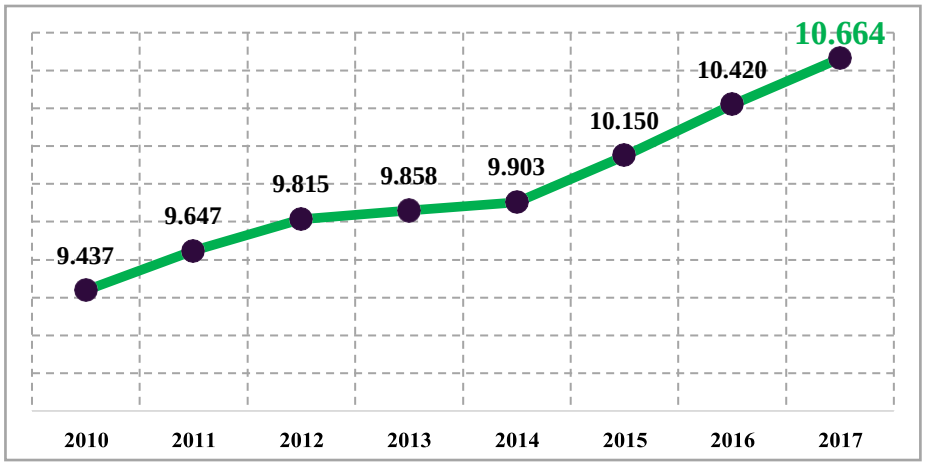
Tabel 2. 5 Harapan Lama Sekolah (HLS) dan Rata-Rata Lama Sekolah (RLS) Menurut Provinsi di Indonesia Tahun 2017

Provinsi	**HLS (tahun)**	**RLS (tahun)**
Aceh	14,13	8,98
Sumatera Utara	13,10	9,25
Sumatera Barat	13,94	8,72
Riau	13,03	8,76
Jambi	12,87	8,15
Sumatera Selatan	12,35	7,99
Bengkulu	13,57	8,47
Lampung	12,46	7,79
Kep. Bangka Belitung	11,83	7,78
Kepulauan Riau	12,81	9,79
DKI Jakarta	12,86	11,02
Jawa Barat	12,42	8,14
Jawa Tengah	12,57	7,27
D I Yogyakarta	15,42	9,19
Jawa Timur	13,09	7,34
Banten	12,78	8,53
Bali	13,21	8,55
Nusa Tenggara Barat	13,46	6,90
Nusa Tenggara Timur	13,07	7,15
Kalimantan Barat	12,50	7,05
Kalimantan Tengah	12,45	8,29
Kalimantan Selatan	12,46	7,99
Kalimantan Timur	13,49	9,36
Kalimantan Utara	12,79	8,62
Sulawesi Utara	12,66	9,14
Sulawesi Tengah	13,04	8,29
Sulawesi Selatan	13,28	7,95
Sulawesi Tenggara	13,36	8,46
Gorontalo	13,01	7,28
Sulawesi Barat	12,48	7,31
Maluku	13,91	9,38
Maluku Utara	13,56	8,61
Papua Barat	12,47	7,15
Papua	10,54	6,27
INDONESIA	**12,85**	**8,10**

2.1.3.3 Standar Hidup Layak

Dimensi standar hidup layak merupakan representasi dari kesejahteraan. Dimensi ini dicerminkan oleh indikator pengeluaran per kapita yang disesuaikan. Pengeluaran per kapita yang disesuaikan di Indonesia terus mengalami peningkatan dari tahun ke tahun. Selama 2010 hingga 2017, pengeluaran per kapita Indonesia meningkat sebesar Rp1.227.000,00 atau tumbuh 1,76 persen per tahun. Jika pada tahun 2010 pengeluaran per kapita penduduk hanya sekitar 9,41 juta rupiah per tahun, maka pada tahun 2017 sudah mencapai 10,66 juta rupiah per tahun. Hal ini menunjukkan bahwa terus membaiknya kesejahteraan penduduk di Indonesia (Badan Pusat Statistik, 2018).

Tabel 2. 6 Pertumbuhan Harapan Lama Sekolah (HLS) dan Rata-Rata Lama Sekolah (RLS) di Indonesia, 2010-2017



Pengeluaran per kapita yang disesuaikan ditentukan dari nilai pengeluaran per kapita dan paritas daya beli (*Purchasing Power Parity*–PPP). Rata-rata pengeluaran per kapita setahun diperoleh dari Susenas dihitung dari level provinsi hingga level kabupaten/kota. Rata-rata pengeluaran per kapita dibuat konstan/riil dengan tahun dasar 2012=100. Perhitungan Paritas Daya Beli pada metode baru menggunakan 96 komoditas dimana 66 komoditas merupakan makanan dan sisanya (30) merupakan komoditas non makanan. Metode penghitungan paritas daya beli menggunakan Metode Rao (BPS Kabupaten Klaten, 2017).

Tabel 2. 7 Pengeluaran per Kapita yang Disesuaikan Menurut Provinsi di Indonesia Tahun 2017

Provinsi	Pengeluaran (ribu rupiah)
Aceh	8.957
Sumatera Utara	10.036
Sumatera Barat	10.306
Riau	10.677
Jambi	9.880
Sumatera Selatan	10.220
Bengkulu	9.778
Lampung	9.413
Kep. Bangka Belitung	12.066
Kepulauan Riau	13.566
DKI Jakarta	17.707
Jawa Barat	10.285
Jawa Tengah	10.377
D I Yogyakarta	13.521
Jawa Timur	10.973
Banten	11.659
Bali	13.573
Nusa Tenggara Barat	9.877
Nusa Tenggara Timur	7.350
Kalimantan Barat	8.472
Kalimantan Tengah	10.492
Kalimantan Selatan	11.600
Kalimantan Timur	11.612
Kalimantan Utara	8.643
Sulawesi Utara	10.422
Sulawesi Tengah	9.311
Sulawesi Selatan	10.489
Sulawesi Tenggara	9.094
Gorontalo	9.532
Sulawesi Barat	8.736
Maluku	8.433
Maluku Utara	7.792
Papua Barat	7.493
Papua	6.996
INDONESIA	**10.664**

2.2 Analisis *Cluster*

2.2.1 Definisi Analisis *Cluster*

Analisis kelompok atau biasa dikenal sebagai *cluster analysis* adalah salah satu teknik statistik yang bertujuan untuk mengelompokkan objek ke dalam suatu kelompok sedemikian sehingga objek yang berada dalam satu kelompok akan memiliki kesamaan yang tinggi dibandingkan dengan objek yang berada di kelompok lain (Sharma, 1996). Analisis Kelompok dengan metode *K-Means* adalah statistik yang berguna untuk mengelompokkan sejumlah objek dalam jumlah kelompok yang sudah ditentukan di mana karakteristik objek hanya dikelompokkan berdasarkan variabel tertentu, tetapi karakteristik latar belakang objek belum diketahui pasti (Yamin & Kurniawan, 2011). *Cluster* atau 'kluster' dapat diartikan 'kelompok'. Dengan demikian, pada dasarnya analisis *cluster* akan menghasilkan sejumlah kluster (kelompok). Analisis ini diawali dengan pemahaman bahwa sejumlah data tertentu sebenarnya mempunyai kemiripan di antara anggotanya; karena itu, dimungkinkan untuk mengelompokkan anggota-anggota yang mirip atau mempunyai karakteristik yang serupa tersebut dalam satu atau lebih dari satu kluster (Santoso, 2014).

2.2.2 Hal-Hal Pokok Tentang Analisis *Cluster*

Menurut Santoso (2014)**,** ciri-ciri sebuah *cluster* yang baik adalah *cluster* yang mempunyai:

1. Homogenitas (kesamaan) yang tinggi antara anggota dalam satu *cluster* (*within-cluster*). Sebagai contoh, *cluster* "konsumen rumah yang peduli lingkungan" tentu terdiri atas orang-orang yang mengutamakan kebersihan dan kenyamanan lingkungan rumahnya; mereka yang mengutamakan harga rumah murah tentu tidak dapat digabungkan menjadi 'anggota' *cluster* tersebut.
2. Heterogenitas (perbedaan) yang tinggi antara *cluster* yang satu dengan c*luster* lainnya (*between cluster*). Dalam contoh di atas, anggota *cluster* konsumen rumah yang peduli lingkungan tentu mempunyai pendapat yang jelas berbeda dengan anggota-anggota c*luster* konsumen rumah yang mementingkan harga murah.

Berdasarkan 2 hal di atas dapat disimpulkan bahwa sebuah *cluster* yang baik adalah *cluster* yang mempunyai anggota-anggota yang semirip mungkin satu dengan yang lain, namun sangat tidak mirip dengan anggota-anggota *cluster* yang lain. Di sini, 'mirip' diartikan sebagai tingkat kesamaan karakteristik antara dua data.

2.2.3 Proses Dasar dari Analisis *Cluster*

Menurut Santoso (2014), karena proses *clustering* pada dasarnya mencari dan mengelompokkan data yang mirip satu dengan yang lain, maka kriteria 'mirip' (*similarity*) adalah dasar dari metode *clustering*. Proses pengolahan data sehingga sekumpulan data mentah dapat dikelompokkan menjadi satu atau beberapa *cluster* adalah sebagai berikut.

1. Menetapkan ukuran jarak antar-data

   Mengukur kesamaan antar-objek (similarity). Sesuai prinsip dasar *cluster* yang mengelompokkan objek yang mempunyai kemiripan, maka proses pertama adalah mengukur seberapa jauh ada kesamaan abtar-objek. Ada 3 metode yang digunakan:

   a. Mengukur korelasi antara sepasang objek pada beberapa variabel. Cara ini sebenarnya sederhana; jika beberapa datamemang akan 'tergabung' menjadi 1 *cluster*, tentulah di antara data tersebut ada hubungan yang erat, atau disebut berkorelasi 1 dengan yang lain. Metode ini mendasarkan besaran korelasi antara data untuk mengetahui kemiripan data 1 dengan yang lain.

   b. Mengukur jarak (*distance*) antara dua objek. Pengukuran ada bermacam-macam, yang paling populer adalah metode Euclidean Distance. Pada dasarnya, cara ini akan memasukkan sebuah data ke dalam *cluster* tertentu yang mengukur 'jarak' data tersebut dengan pusat *cluster*. Jika data ada dalam jarak yang masih ada dalam batas tertentu, data tersebut dapat dimasukkan pada *cluster* tersebut.

c. Mengukur asosiasi antar-objek. Pada dasarnya, cara ini akan mengasosiasikan sebuah data dengan *Cluster* tertentu; dalam praktek, cara ini tidak sepopuler kedua cara sebelumnya.

2. Melakukan proses standardisasi data jika diperlukan

Setelah cara mengukur jarak ditetapkan, yang juga perlu diperhatikan adalah apakah satuan data mempunyai perbedaan yang besar. Sebagai contoh, jika variabel Penghasilan mempunyai satuan juta (000.000), sedangkan usia seseorang hanya mempunyai satuan puluhan (00), maka perbedaan yang mencolok ini akan membuat perhitungan jarak (*distance*) menjadi tidak valid. Jika data memang mempunyai satuan yang berbeda secara signifikan, pada data harus dilakukan proses standardisasi dengan mengubah data yang ada ke Z-Score. Proses standardisasi menjadikan dua data dengan perbedaan satuan yang lebar akan otomatis menjadi menyempit.

3. Melakukan proses *Cluster*ing

Setelah data yang dianggap mempunyai satuan yang sangat berbeda diseragamkan, dan metode *Cluster* ditentukan (misal dipilih Eucledian), langkah selanjutnya adalah membuat *Cluster*. Proses inti dari *Clustering* adalah pengelompokkan data, yang bisa dilakukan dengan 2 metode:

a. HIERARCHICAL METHOD

Metode ini memulai pengelompokkan dengan 2 atau lebih objek yang mempunyai kesamaan paling dekat. Kemudian proses diteruskan ke objek lain yang mempunyai kedekatan kedua. Demikian seterusnya sehingga *Cluster* akan jelas antar-objek, dari yang paling mirip sampai paling tidak mirip. Secara logika semua objek pada akhirnya hanya akan membentuk sebuah *Cluster*. Dendogram biasanya digunakan untuk membantu memperjelas proses hierarki tersebut.

b. NON-HIERARCHICAL METHOD

Berbeda dengan metode hierarki, metode ini justru dimulai dengan menentukan terlebih dahulu jumlah *Cluster* yang diinginkan (2 *Cluster*, 3 *Cluster* atau yang lain). Setelah jumlah *Cluster* diketahui, baru proses

Cluster dilakukan tanpa proses hierarki. Metode ini biasa disebut dengan *K-Means Cluster*.

4. Melakukan penamaan *Cluster-Cluster* yang terbentuk

   Setelah sejumlah *cluster* terbentuk, entah dengan metode hierarki atau non-hierarki, langkah selanjutnya adalah melakukan interpretasi terhadap *cluster* yang terbentuk, yang pada intinya memberi nama spesifik untuk menggambarkan isi *cluster* tersebut. Misal kelompok konsumen yang memerhatikan lingkungan sekitar sebelum membeli sebuah rumah bisa dinamai *cluster* LINGKUNGAN.

5. Melakukan validasi dan profiling kluster

   Kluster yang terbentuk kemudian diuji apakah hasil tersebut valid. Kemudian dilakukan proses profiling untuk menjelaskan karakteristik setiap *cluster* berdasar profil tertentu (seperti usia konsumen pembeli rumah, tingkat penghasilannya dan sebagainya). Dari data profiling tersebut bisa dilakukan analisis lanjutan seperti Analisis Diskriminan.

2.2.4 Metode Analisis *Cluster*

Analisis *Cluster* terdapat dua metode yaitu metode hierarki dan metode non-hierarki. Keduanya memiliki kelebihan dan kekurangan masing-masing. Kelebihan metode hierarki yaitu cepat dalam proses pengolahan sehingga menghemat waktu, namun kekurangannya yaitu metode ini dapat menimbulkan kesalahan. Selain itu tidak baik diterapkan untuk menganalisis sampel dengan ukuran besar. Metode non-hierarki memiliki kelebihan yang lebih baik dibandingkan metode hierarki. Hasilnya memiliki sedikit kelemahan pada data *outlier*. Ukuran jarak yang digunakan dan termasuk variabel tak relevan atau variabel yang tidak tepat. Keuntungannya adalah hanya dengan menggunakan titik bakal non *random*, penggunaan metode non-hierarki untuk titik bakal *random* secara nyata lebih buruk dari pada metode hierarki. Penentuan metode ini yang akan dipakai tergantung kepada peneliti dan konteks penelitian dengan tidak mengabaikan substansi, teori dan konsep yang berlaku. Alternatif lain adalah dengan mengkombinasikan kedua

metode ini. Pertama gunakan metode hierarki kemudian dilanjutkan dengan metode non-hierarki (Supranto, 2004).

1. Metode Hierarki

   Metode ini memulai pengelompokkan dengan dua atau lebih objek yang mempunyai kesamaan paling dekat. Kemudian proses diteruskan ke objek lain yang mempunyai kedekatan kedua. Demikian seterusnya sehingga *Cluster* akan membentuk semacam 'pohon' dimana ada hierarki (tingkatan) yang jelas antar-objek, dari yang paling mirip sampai paling tidak mirip. Secara logika semua objek pada akhirnya hanya akan membentuk sebuah *Cluster*. *Dendogram* biasanya digunakan untuk membantu memperjelas proses hierarki tersebut (Santoso, Statistik Multivariat Konsep dan Aplikasi dengan SPSS, 2014).

   Menurut Utomo (2017), Andaikan dij merupakan ukuran ketidakmiripan antara *Cluster* ke-*i* dengan *Cluster* ke-*j* dan $dk(i,j)$ merupakan ukuran ketidakmiripan antara *Cluster* ke-*k* dengan klaster (*i*,*j*) yang merupakan penggabungan antara *Cluster* ke-*i* dengan klaster ke-*j*. Berikut ini diberikan ukuran ketidakmiripan antar *Cluster*.

   a. Pautan Tunggal (*Single Linkage/Nearest Neighbour Method*)

      Metode ini akan mengelompokkan dua objek yang mempunyai jarak terdekat terlebih dahulu (Santoso, Statistik Multivariat Konsep dan Aplikasi dengan SPSS, 2014). Ukuran ketidakmiripan yang digunakan adalah

      $$d_{k(i,j)} = \min\left(d_{ki}, d_{kj}\right)$$

   b. Pautan Lengkap (*Complete Linkage/Furthest Neighbour Method*)

      Metode ini akan mengelompokkan dua objek yang mempunyai jarak terjauh terlebih dahulu (Santoso, 2014). Ukuran ketidakmiripan yang digunakan adalah

      $$d_{k(i,j)} = \text{maks}(d_{ki}, d_{kj})$$

c. Pautan Rataan (*Avarage Linkage/Between Neighbour Method*)

Metode ini akan mengelompokkan objek berdasar jarak rata-rata yang didapat dengan melakukan rata-rata semua jarak antar objek terlebih dahulu (Santoso, 2014). Ukuran ketidakmiripan yang digunakan adalah

$$d_{k(i,j)} = \frac{n_i}{n_i + n_j} d_{ki} + \frac{n_j}{n_i + n_j} d_{kj}$$

Dengan n_i adalah jumlah item pada *Cluster* i.

d. Metode *Centroid*

Metode ini akan mengelompokkan objek berdasar jarak rata-rata yang didapat dengan melakukan rata-rata semua jarak antar objek terlebih dahulu (Santoso, 2014). Ukuran ketidakmiripan yang digunakan adalah

$$d_{k(i,j)} = \frac{n_i}{n_i + n_j} d_{ki} + \frac{n_j}{n_i + n_j} d_{kj} - \frac{n_i + n_j}{(n_i + n_j)^2} d_{ij}$$

2. Metode Non- Hierarki

Berbeda dengan metode hierarki, metode ini justru dimulai dengan menentukan terlebih dahulu jumlah *Cluster* yang diinginkan (dua *Cluster*, tiga *Cluster* atau yang lain). Setelah jumlah *Cluster* diketahui, baru proses *Cluster* dilakukan tanpa mengikuti proses hierarki. Metode ini biasa disebut dengan *K-Means Cluster* (Santoso, Statistik Multivariat Konsep dan Aplikasi dengan SPSS, 2014). Ada tiga prosedur dalam metode non-hierarki yaitu :

a. Sequential Threshold

Proses clusteing dengan menggunakan metode *Sequential Threshold* dimulai dengan memilih sebuah klaster *seed* dan kemudian menggabungkan setiap objek yang ada dalam jarak yang telah ditentukan sebelumnya. Klaster tersebut akan disebut sebagai klaster pertama. Setelah klaster pertama terbentuk, maka klaster *seed* kedua digunakan dan kemudian objek-objek yang mempunyai jarak terdekat akan digabungkan. Jika telah selesai, maka klaster berikutnya akan dibentuk dengan cara yang sama.

Metode ini disebut *Sequential Threshold* karena proses *Clustering* dilakukan berurutan dari klaster pertama, kedua, ketiga dan berikutnya. Metode ini tidak memperbolehkan suatu objek pindah ke klaster lain.

b. Parallel Threshold

Proses *Clustering* dengan menggunakan metode *Parallel Threshold* dimulai dengan memiliki klaster *seed* yang akan dijadikan patokan pembuatan klaster. Setiap objek akan diukur tehadap klaster *seed* tersebut. Sebuah objek akan masuk ke suatu klaster jika mempunyai nilai jarak terhadap suatu klaster *seed* lebih dekat daripada klaster *seed* yang lain. Langkah ini mengakibatkan penentuan klaster tidak berurutan.

Metode ini disebut *Parallel Threshold* karena proses *Clustering* dilakukan tidak berurutan. Metode ini tidak memperbolehkan suatu objek yang sudah menjadi anggota suatu klaster berpindah ke klaster lain.

c. Optimalisasi

Proses *Cluster*ing dengan menggunakan metode ini mirip dengan *Sequential Threshold* atau *Parallel Threshold Clustering*, perbedaannya adalah diperbolehkannya suatu objek berpindah dari suatu klaster ke klaster lain. Ini tejadi ketika jarak ke suatu klaster ternyata lebih dekat daripada jarak objek tersebut ke klaster yang sekarang.

Proses optimalisasi dilakukan dengan menghitung ulang setiap objek terhadap semua *centroid* klaster yang ada. Jika objek tersebut mempunyai jarak yang lebih kecil terhadap suatu *centroid* klaster lain daripada *centroid* klaster sekarang, maka objek tersebut di relokasi ke klaster terdekat tersebut.

2.2.5 *K-Means Cluster*

K-Means Cluster Analysis merupakan salah satu metode analisis *Cluster* non hirarki yang berusaha untuk mempartisi objek yang ada ke dalam satu atau lebih *Cluster* atau kelompok objek berdasarkan karakteristiknya, sehingga objek yang mempunyai karakteristik yang sama dikelompokkan dalam satu *Cluster* yang sama dan objek yang mempunyai karakteristik yang berbeda dikelompokkan ke dalam *Cluster* yang lain. Tujuan pengelompokan adalah untuk meminimalkan *objective function* yang di set dalam proses *Clustering*, yang pada dasarnya berusaha untuk meminimalkan variasi dalam satu *Cluster* dan memaksimalkan variasi antar *Cluster*.

Metode *Cluster* ini meliputi *sequential threshold, pararel threshold dan optimizing threshold. Sequential threshold* melakukan pengelompokan dengan terlebih dahulu memilih satu objek dasar yang akan dijadikan nilai awal *Cluster*, kemudian semua *Cluster* yang ada dalam jarak terdekat dengan *Cluster* ini akan bergabung, lalu dipilih *Cluster* kedua dan semua objek yang mempunyai kemiripan dengan *Cluster* ini akan digabungkan, demikian seterusnya sehingga terbentuk beberapa *Cluster* dengan keseluruhan objek yang terdapat di dalamnya.

Jika diberikan sekumpulan objek $X = (x_1, x_2, , . . , x_n)$ maka algoritma *K-Means Cluster Analysis* akan mempartisi *X* dalam *k* buah *Cluster*, setiap *Cluster* memiliki *centroid* dari objek-objek dalam *Cluster* tersebut. Pada tahap awal algoritma *K-Means Cluster Analysis* dipilih secara acak *k* buah objek sebagai *centroid*, kemudian jarak antara objek dengan *centroid* dihitung dengan menggunakan jarak *euclidian*, objek ditempatkan dalam *Cluster* yang terdekat dihitung dari titik tengah *Cluster*. *Centroid* baru ditetapkan jika semua objek sudah ditempatkan dalam *Cluster* terdekat. Proses penentuan *centroid* dan penempatan objek dalam *Cluster* diulangi sampai nilai *centroid* konvergen (*centroid* dari semua *Cluster* tidak berubah lagi).

Menurut Ediyanto (2013), Secara umum metode *K-Means Cluster Analysis* menggunakan algoritma sebagai berikut:

1. Tentukan *k* sebagai jumlah *Cluster* yang di bentuk.

   Untuk menentukan banyaknya *Cluster k* dilakukan dengan beberapa pertimbangan seperti pertimbangan teoritis dan konseptual yang mungkin diusulkan untuk menentukan berapa banyak *Cluster*.

2. Bangkitkan *k Centroid* (titik pusat *Cluster*) awal secara *random.*

   Penentuan *centroid* awal dilakukan secara *random*/acak dari objek-objek yang tersedia sebanyak *k Cluster*, kemudian untuk menghitung *centroid Cluster* ke-*i* berikutnya, digunakan rumus sebagai berikut :

$$v = \frac{\sum_{i=1}^{n} x_i}{n} \quad : i = 1,2,3,\ldots,n$$

   Dimana :

   v : *centroid* pada *Cluster*

   x_i : objek ke-*i*

   n : banyaknya objek/jumlah objek yang menjadi anggota *Cluster*

3. Hitung jarak setiap objek ke masing-masing *centroid* dari masing-masing *Cluster*.

   Untuk menghitung jarak antara objek dengan *centroid* penulis menggunakan *Euclidian Distance.*

$$d(x,y) - \left||x-y|\right| - \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2} \quad ; i = 1,2,3,\ldots,n$$

   Dimana :

   x_i : objek *x* ke-i

   y_i : daya *y* ke-i

   n : banyaknya objek

4. Alokasikan masing-masing objek ke dalam *centroid* yang paling terdekat.

   Untuk melakukan pengalokasian objek kedalam masing-masing *Cluster* pada saat iterasi secara umum dapat dilakukan dengan dua cara yaitu dengan *hard K-Means*, dimana secara tegas setiap objek dinyatakan sebagai anggota *Cluster* dengan mengukur jarak kedekatan sifatnya terhadap titik pusat *Cluster* tersebut, cara lain dapat dilakukan dengan *fuzzy C-Means.*

5. Lakukan iterasi, kemudian tentukan posisi *centroid* baru
6. Ulangi langkah 3 jika posisi *centroid* baru tidak sama.

   Pengecekan *konvergensi* dilakukan dengan membandingkan matriks *group assignment* pada iterasi sebelumnya dengan matrik *group assignment* pada iterasi yang sedang berjalan. Jika hasilnya sama maka *algoritma K-Means Cluster analysis* sudah *konvergen*, tetapi jika berbeda maka belum *konvergen* sehingga perlu dilakukan iterasi berikutnya.

2.3 Analisis Diskriminan

2.3.1 Definisi Analisis Diskriminan

Menurut Santoso (2014), Analisis Diskriminan adalah teknik multivariat yang termasuk pada *Dependence Method*, dengan ciri adanya variabel dependen dan independen. Dengan demikian, ada variabel yang hasilnya tergantung pada data variabel independen. Ciri khusus analisis diskriminan adalah data variabel dependen harus berupa data kategori, sedangkan data untuk variabel independen justru berupa data rasio.

Santoso (2014) juga menjelaskan bahwa, secara teknis, analisis diskriminan mirip dengan analisis regresi, karena keduanya mempunyai variabel dependen dan variabel independen dalam modelnya. Hanya pada analisis regresi (sederhana maupun berganda), vari abel dependen harus data rasio; sedangkan pada analisis diskriminan, jenis data untuk variabel dependen harus kategori.

2.3.2 Kegunaan Analisis Diskriminan

Menurut Santoso (2014), karena mempunyai model yang sama, secara dasar kegunaan, analisis diskriminan sama dengan analisis regresi. Dengan demikian, kegunaan utama dari analisis diskriminan ada 2, yaitu:

1. Kemampuan memprediksi terjadinya variabel dependen dengan masukan data variabel independen.
2. Kemampuan memilih mana variabel yang secara nyata mempengaruhi variabel dependen dan mana yang tidak.

2.3.3 Tujuan Analisis Diskriminan

Menurut Santoso (2014), karena bentuk multivariat dari Analisis Diskriminan adalah Dependence, maka variabel Dependen adalah variabel yang menjadi dasar analisis diskriminan. Variabel Dependen bisa berupa kode grup 1 atau grup 2 atau lainnya, dengan tujuan diskriminan secara umum adalah:

1. Ingin mengetahui apakah ada perbedaan yang jelas antar-grup pada variabel dependen? Atau bisa dikatakan spskah ada perbedaan antara grup 1 dengan anggota grup 2?
2. Jika ada perbedaan, variabel independen manakah pada fungsi diskriminan yang membuat perbedaan tersebut?
3. Membuat Fungsi atau Model Diskriminan, yang pada dasarnya mirip dengan persamaan regresi.
4. Melakukan klasifikasi terhadap objek (dalam terminologi SPSS disebut baris), apakah suatu objek (bisa nama orang, nama tumbuhan, benda, atau yang lainnya) termasuk pada grup 1 atau grup 2, atau lainnya.

2.3.4 Jumlah Sampel Diskriminan

Asra et al. (2017) menjelaskan bawha pada dasarnya tidak terdapat aturan tentang jumlah sampel minimum yang digunakan dalam analisis diskriminan. Namun beberapa referensi menyebutkan bahwa semakin besar ukuran sampel, makin stabil nilai estimasi parameternya. Rencher (2002) mencontohkan bahwa hasil estimasi akan lebih stabil jika diperoleh dari 50 observasi dengan 2 variabel dibandingkan diperoleh dari 50 observasi dengan 20 variabel.

Santoso (2014) juga menyebutkan bahwa secara pasti tidak ada jumlah sampel yang ideal pada analisis diskriminan. Pedoman yang bersifat umum menyatakan untuk setiap variabel independen sebaiknya ada 5-20 data (sampel). Jadi, apabila ada 6 variabel independen, seharusnya minimal ada 6x5 = 30 sampel. Secara terminologi SPSS, jika ada 6 variabel independen, sebaiknya ada 30 baris data.

2.3.5 Asumsi pada Analisis Diskriminan

Sebelum melakukan analisis diskriminan harus dipenuhi beberapa asumsi analisis diskriminan dan menguji variabel apa saja yang layak digunakan untuk analisis diskriminan (Rizkiana & Hendikawati, 2015).

Dijelaskan pula oleh Santoso (2014), asumsi penting yang harus dipenuhi agar model diskriminan bisa digunakan adalah:

1. *Multivariate Normality*, atau variabel independen seharusnya berdistribusi normal. Jika data tidak berdistribusi normal, hal ini akan menyebabkan masalah pada ketepatan fungsi (model) diskriminan.
2. Matriks kovarians dari semua variabel independen seharusnya sama (*equal*).
3. Tidak ada korelasi antar-variabel independen. Jika dua variabel independen mempunyai korelasi yang kuat, dikatakan terjadi multikolinieritas.
4. Tidak adanya data yang sangat ekstrem (outlier) pada variabel independen. Jika ada data outlier yang tetap diproses, hal ini bisa berakibat berkurangnya ketepatan klasifikasi dari fungsi diskriminan.

2.3.6 Proses Dasar Analisis Diskriminan

Proses dasar analisis diskriminan menurut Santoso (2014) yaitu:

1. Memisah variabel-variabel menjadi Variabel Dependen dan Variabel Independen
2. Menentukan metode untuk membuat Fungsi Diskriminan. Pada prinsipnya ada 2 metode dasar untuk itu, yakni:
   a. SIMULTANEOUS ESTIMATION, dimana semua variabel dimasukkan secara bersama-sama kemudian dilakukan proses diskriminan.
   b. STEP-WISE ESTIMATION, dimana variabel dimasukkan satu per satu ke dalam model diskriminan. Pada proses ini, tentu ada variabel yang tetap ada pada model, dan ada kemungkinan satu atau lebih variabel independen yang ‘dibuang’ dari model.
3. Menguji signifikansi dari fungsi diskriminan yang telah terbentuk, menggunakan Wilk’s Lambda, Pilai, F test dan lainnya.

4. Menguji ketepatan klasifikasi dari fungsi diskriminan, termasuk mengetahui ketepatan klasifikasi secara individual dengan Casewise Diagnostics.
5. Melakukan interpretasi terhadap fungsi diskriminan tersebut.
6. Melakukan uji validasi fungsi diskriminan.

Langkah-langkah dalam melakukan analisis diskriminan menurut Supranto (2004) yaitu:

1. Merumuskan masalah
   a. Memasukkan objek ke dalam kategori
   b. Menentukan variabel bebas dan terikat
2. Membuat perkiraan koefisien diskriminan

   Terdapat dua metode yaitu *simultaneous method* (variabel independen yang ada dipertimbangkan secara bersama) dan *stepwise method* (menyeleksi otomotis untuk memilih variabel terbaik).
3. Mengestimasi koefisien fungsi diskriminan
4. Melakukan uji rata-rata kelompok
5. Melakukan uji signifikan
6. Melakukan ketepatan fungsi diskriminan
7. Menginterpretasikan koefisien fungsi diskriminan
   d. Menghitung nilai diskriminan

      Rumus fungsi persamaan diskriminan yaitu:

      $$Z = a + b_1X_1 + b_2X_2 + b_3X_3 + \ldots + b_nX_n$$

      Keterangan:

      Z = Nilai diskriminan

      a = konstanta

      $b_{1\ldots n}$ = koefisien

      $X_{1\ldots n}$ = variabel bebas

e. Menentukan titik *cut off*

$$\boxed{Z_{CE} = \frac{Z_A + Z_B}{2}}$$

Keterangan:

Z_{CE} = nilai kritis

Z_A = rata-rata provinsi dengan nilai IPM rendah

Z_B = rata-rata provinsi dengan nilai IPM tinggi

Jika pengelompokkan perusahaan mempunyai ukuran sampel yang tidak sama maka menggunakan rumus sebagai berikut.

$$\boxed{Z_{CE} = \frac{N_A Z_A + N_B Z_B}{N_A N_B}}$$

Keterangan:

Z_{CE} = nilai kritis

Z_A = rata-rata provinsi dengan nilai IPM rendah

Z_B = rata-rata provinsi dengan nilai IPM tinggi

N_A = jumlah provinsi dengan nilai IPM rendah

N_B = jumlah provinsi dengan nilai IPM tinggi

8. Uji validitas analisis diskriminan

BAB V
PENUTUP

5.1 Simpulan

Berdasarkan hasil analisis dan pembahasan yang telah dipaparkan pada bab sebelumnya, diperoleh simpulan sebagai berikut.

1. Hasil analisis pengelompokkan menggunakan metode *K-Means Cluster* berdasarkan indikator penyusun Indeks Pembangunan Manusia tahun 2017 diperoleh 2 kelompok (*cluster*) yang terbentuk. Karakteristik untuk *cluster* 1 yaitu berisi provinsi dengan umur harapan hidup (UHH) yang relatif rendah, harapan lama sekolah (HLS) yang relatif rendah, rata-rata lama sekolah (RLS) yang relatif rendah, dan pengeluaran disesuaikan yang rendah juga, sedangkan pada cluster 2 berisi provinsi dengan umur harapan hidup (UHH) relatif tinggi, harapan lama sekolah (HLS) yang relatif tinggi, rata-rata lama sekolah (RLS) yang relatif tinggi, dan pengeluaran disesuaikan yang relatif tinggi juga.
2. Model diskriminan dua faktor yang terbentuk dari hasil pengelompokan provinsi di Indonesia menggunakan metode *K-Means Cluster* berdasarkan indikator penyusun Indeks Pembangunan Manusia tahun 2017 adalah $Y = -11{,}642 + 1{,}409\, RLS$. Diketahui bahwa faktor yang paling berpengaruh dari penyusun nilai indeks pembangunan manusia yaitu faktor dari rata-rata lama sekolah di setiap provinsi di Indonesia.
3. Ketepatan hasil dalam mengklasifikasi kasus dari model diskriminan yang terbentuk adalah sebesar 91.20% maka dapat disimpulkan bahwa tingkat keakuratan hasil pengelompokan *K-Means Cluster* tergolong tinggi karena ketepatan prediksi mendekati angka 100%.

5.2 Saran

1. Pendidikan adalah hal penting untuk memajukan pembangunan manusia dalam suatu daerah atau negara. Maka dari itu, saran dari penulis untuk pemerintah adalah memperhatikan pendidikan di seluruh Indonesia ini, terutama di Indonesia bagian timur dengan cara memberikan fasilitas pendidikan berupa saran dan prasarana yang mencukupi serta menjamin kesejahteraan tenaga pendidik yang berada disana. Sehingga nanti harapannya tercipta SDM unggul dan Indonesia maju seperti yang diharapkan di HUT ke-74 Kemerdekaan Republik Indonesia.
2. Model diskriminan pada penelitian ini dapat digunakan untuk meramalkan indeks pembangunan manusia di Indonesia
3. Hasil penelitian ini dapat digunakan sebagai salah satu bahan untuk penelitian lanjutan, mungkin bisa ditambahkan untuk faktor-faktor yang lain.

DAFTAR PUSTAKA

Asra, A., Utomo, A., Asikin, M., & Pusponegoro, N. (2017). *ANALISIS MULTIVARIABEL: SUATU PENGANTAR.* Bogor: Penerbit IN MEDIA.

Badan Pusat Statistik. 2018. *Indeks Pembangunan Manusia 2017*. Tersedia di https://www.bps.go.id/publication/2018/08/27/75895ba5f85eb0a77f75c339/indeks-pembangunan-manusia-2017.html. Diakses Jum'at 9 November 2018

Badan Pusat Statistik Kabupaten Klaten. 2017. *Indeks Pembangunan Manusia Kabupaten Klaten 2016*. Tersedia di https://klatenkab.bps.go.id/publication/2018/01/24/31948f716a52b54b42fc37f4/indeks-pembangunan-manusia-kabupaten-klaten-2016.html. Diakses Jum'at 9 November 2018

Ding, C., & He, X. (2004). K-Means Clustering via Principal Component Analysis. *International Conference on Machine Learning.* Canada: Banff.

Dris, J., & Tasari. (2011). *Matematika Jilid 1 untuk SMP dan MTs Kelas VII.* Jakarta: Pusat Kurikulum dan Perbukuan Kementrian Pendidikan Nasional.

Ediyanto, Mara, M. N., & Satyahade, N. (2013). PENGKLASIFIKASIAN KARAKTERISTIK DENGAN METODE K-MEANS CLUSTER ANALYSIS. *Buletin ilmiah Mat, Stat, dan Terapannya (Bimaster)*, 1.

Faisyal. (2014). MENGATASI MULTIKOLINEARITAS MENGGUNAKAN PRINCIPAL COMPONENT ANALYSIS (PCA).

Fathia, A. N., Rahmawati, R., & Tarno. (2016). Analisis Klaster Kecamatan di Kabuupaten semarang Berdasarkan Potensi Desa Menggunakan Metode Ward an Single Linkage. *Gaussian, 5*, 801-810.

Gaspersz, V. (1989). *Statistika.* Bandung: Armico.

Hair, J. F., Tatham, R. L., & Anderson, R. E. (2006). *Multivariate Data Analysis* (5 ed.). Jakarta: Gramedia Pstaka Utama.

Juaeni, I. (2014). Dampak Penerapan Principal Component Analysis (PCA) dalam Clstering Curah Hujan di Pulau Jawa, Bali, dan Lombok. *Lapan*.

Melinda. (2006). Analisis Pengelompokan Objek dengan metode Single Linkage Clustering dan Diskriminan Linear untuk Kasus Dua Kelompok.

Mustakim. (2017). Centroid K-Means Clustering Optimiztion Using Eigenvector Principal Component Analysis. *Theoritical and Applied Information Technology*.

Nuharini, D., & Wahyuni, T. (2008). *Matematika Konsep dan Aplikasinya Untuk Kelas VII SMP/MTs.* Jakarta: Pusat Perbukuan Departemen Pendidikan Nasional.

Nuharini, D., & Wahyuni, T. (2008). *Matematika Konsep dan Aplikasinya Untuk SMP/MTs Kelas VIII.* Jakarta: Pusat Perbukuan Departemen Pendidikan Nasional.

Pratowo, N. I. (2013). ANALISIS FAKTOR-FAKTOR YANG BERPENGARUH TERHADAP INDEKS. *Jurnal Studi Ekonomi Indonesia*.

Putra, S. F., Pradina, R., & Hafidz, I. (2016). Feature Selection pada Dataset Faktor Kesiapan Bencana pada Provinsi di Indonesia Menggunakan Metode PCA. *Teknik ITS, 5*.

Putri, R., & Widodo, E. (2017). Analisis Klaster Hierarki untuk Pengelompokan Kabupaten/Kota Di Jawa Tengah Berdasarkan Indikator IPM tahun 2015. Yogyakarta: Universitas Islam Indonesia.

Rencher, A. (2002). *Methods of Multivariate Analysis.* Second Edition. Brigham Young University.

Rizkiana, A., & Hendikawati, P. (2015). KLASIFIKASI TINGKAT KEPUASAN PELANGGAN DENGAN ANALISIS. *Jurnal MIPA*, 89-100.

Santoso, S. (2014). *Menguasai Statistik Multivariat.* Jakarta: PT Elex Media Komputindo.

Santoso, S. (2014). *Statistik Multivariat Konsep dan Aplikasi dengan SPSS.* Jakarta: Elex Media Komputindo.

Sharma, S. (1996). *Applied Multivariate Techniques.* New York: John Wiley and Sons, Inc.

Sugiyono. (2013). *Metode Penelitian Kuantitatif, Kualitatif dan R&D.* Bandung: ALFABETA.

Supranto, J. (2004). *Analisis Multivariat Arti & Interpretasi.* Jakarta: PT. RINEKA CIPTA.

Thiagarajan, S., Semmel, D. S., & Semmel, M. I. (1974). *Instructional Development for Training Teachers of.* Minnesota: Indiana University.

Utomo, C. A. (2017). *Analisis Multivariat Clustering K-Means untuk APBD Kecamatan-Kecamatan di Kota Semarang.* Semarang: Universitas Negeri Semarang.

Yamin, S., & Kurniawan, H. (2011). *Generasi Baru Mengolah Data Penelitian dengan Partial Least Square Path Modeling : Aplikasi dengan Software XLSTAT, SmartPLS, dan Visual PLS.* Jakarta: Salemba Infotek.