ANALISIS LOYALITAS PELANGGAN AEGIS PADA PEMBELIAN PAKET COACHING MENGGUNAKAN ALGORITMA BISECTING K-MEANS

SKRIPSI

UIN SUNAN AMPEL
S U R A B A Y A

Disusun oleh:

CELVIN GAZINDA

H76216056

PROGRAM STUDI SISTEM INFORMASI

FAKULTAS SAINS DAN TEKNOLOGI

UNIVERSITAS ISLAM NEGERI SUNAN AMPEL

SURABAYA

2021

PERNYATAAN KEASLIAN

Saya yang bertanda tangan di bawah ini,

Nama : Celvin Gazinda

NIM : H76216056

Program Studi : Sistem Informasi

Angkatan : 2016

Menyatakan bahwa saya tidak melakukan plagiat dalam penulisan skripsi saya yang berjudul : "ANALISIS LOYALITAS PELANGGAN AEGIS PADA PEMBELIAN PAKET COACHING MENGGUNAKAN ALGORITMA BISECTING K-MEANS". Apabila suatu saat nanti terbukti bahwa saya melakukan Tindakan plagiat, maka saya bersedia menerima sanksi yang telah ditetapkan.

Demikian pernyataan keaslian ini saya buat dengan sebenar-benarnya.

Surabaya, 22 Juni 2021

Yang menyatakan,

(Celvin Gazinda)

NIM. H76216056

LEMBAR PERSETUJUAN PEMBIMBING

Skripsi oleh

JUDUL : ANALISIS LOYALITAS PELANGGAN AEGIS PADA PEMBELIAN PAKET COACHING MENGGUNAKAN ALGORITMA BISECTING K-MEANS

NAMA : CELVIN GAZINDA

NIM : H76216056

Ini telah diperiksa dan disetujui untuk diujikan.

Surabaya, 5 Mei 2021

Menyetujui,

Dosen Pembimbing 1	Dosen Pembimbing 2
Mujib Ridwan, S. Kom., M.T	Ahmad Yusuf, M. Kom
NIP. 198604272014031004	NIP. 199001202014031003

PENGESAHAN TIM PENGUJI SKRIPSI

Skripsi oleh Celvin Gazinda ini telah dipertahankan di depan tim penguji skripsi di Surabaya, 15 Juni 2021

Mengesahkan,

Dewan Penguji

Penguji I	Penguji II
Mujib Ridwan, S. Kom., M.T	Ahmad Yusuf, M. Kom
NIP. 198604272014031004	NIP. 199001202014031003
Penguji III	Penguji IV
Nurissaidah Ulinnuha, M. Kom	Khalid, M. Kom
NIP. 199011022014032004	NIP. 197906092014031002

Mengetahui,

Dekan Fakultas Sains dan Teknologi

UIN Sunan Ampel Surabaya

Dr. Evi Fatimatur Rusydiyah, M.Ag

NIP. 197312272005012003

KEMENTERIAN AGAMA
UNIVERSITAS ISLAM NEGERI SUNAN AMPEL SURABAYA
PERPUSTAKAAN
Jl. Jend. A. Yani 117 Surabaya 60237 Telp. 031-8431972 Fax.031-8413300
E-Mail: perpus@uinsby.ac.id

LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI
KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademika UIN Sunan Ampel Surabaya, yang bertanda tangan di bawah ini, saya:

Nama : CELVIN GAZINDA
NIM : H76216056
Fakultas/Jurusan : FAKULTAS SAINS DAN TEKNOLOGI / SISTEM INFORMASI
E-mail address : celvingazinda54@gmail.com / h76216056@uinsby.ac.id

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Perpustakaan UIN Sunan Ampel Surabaya, Hak Bebas Royalti Non-Eksklusif atas karya ilmiah :
■ Sekripsi ☐ Tesis ☐ Desertasi ☐ Lain-lain (.....)
yang berjudul :
ANALISIS LOYALITAS PELANGGAN AEGIS PADA PEMBELIAN PAKET COACHING MENGGUNAKAN ALGORITMA BISECTING K-MEANS

beserta perangkat yang diperlukan (bila ada). Dengan Hak Bebas Royalti Non-Ekslusif ini Perpustakaan UIN Sunan Ampel Surabaya berhak menyimpan, mengalih-media/format-kan, mengelolanya dalam bentuk pangkalan data (database), mendistribusikannya, dan menampilkan/mempublikasikannya di Internet atau media lain secara ***fulltext*** untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan atau penerbit yang bersangkutan.

Saya bersedia untuk menanggung secara pribadi, tanpa melibatkan pihak Perpustakaan UIN Sunan Ampel Surabaya, segala bentuk tuntutan hukum yang timbul atas pelanggaran Hak Cipta dalam karya ilmiah saya ini.

Demikian pernyataan ini yang saya buat dengan sebenarnya.

Surabaya, 22 Juni 2021

Penulis

(CELVIN GAZINDA)
nama terang dan tanda tangan

ABSTRAK

ANALISIS LOYALITAS PELANGGAN AEGIS PADA PEMBELIAN PAKET COACHING MENGGUNAKAN ALGORITMA BISECTING K-MEANS

Oleh:

Celvin Gazinda

Teknologi selalu berkembang setiap tahun untuk menunjang kebutuhan sehari-hari manusia. Pada era digital ini banyak yang telah terbantu karena kemajuan teknologi, terutama pada kegiatan *electronic sport* (eSport*)*. Hal ini membuat AEGIS ikut serta membantu para pemain amatir *Dota 2* untuk belajar menekuni *gamenya* agar mampu menjadi pemain professional nantinya. Namun, masalah yang sedang dihadapi oleh AEGIS sekarang adalah jumlah transaksi mengalami penurunan. Berdasarkan masalah tersebut maka dilakukanlah penelitian untuk mencari faktor yang mempengaruhi loyalitas pelanggan AEGIS menggunakan algoritma Bisecting K-Means dengan *machine learning python*. Kemudian, melakukan validasi *cluster* dengan menggunakan metode *Silhouette Coefficient* untuk melihat seberapa kuat kualitas *cluster* tersebut. Terakhir dilakukan analisa *Customer Relationship Management* (CRM) untuk memberikan rekomendasi berupa diskon, melakukan promosi untuk pelanggan baru atau rekomendasi lainnya. Hasil dari penerapan algoritma *Bisecting K-Means* pada data transaksi AEGIS memiliki *score* sebesar 0,81 pada 2 *cluster* dan 0,45 pada 3 cluster. Berdasarkan hasil *score* tersebut, 2 *cluster* dapat dikatakan sebagai *cluster* kuat. Sedangkan, 3 *cluster* dikatakan sebagai *cluster* lemah karena masih ada data yang tidak berada pada *cluster* sebenarnya. Setelah dilakukan analisa, *cluster* 2 termasuk kategori pelanggan yang loyal, sehingga rekomendasi yang diberikan berupa diskon dan kartu member agar tetap menjaga loyalitas pelanggan melakukan pendaftaran.

Kata Kunci : ***Dota 2, machine learning, python, Silhouette Coefficient, CRM***

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

ABSTRACT

ANALISIS LOYALITAS PELANGGAN AEGIS PADA PEMBELIAN PAKET COACHING MENGGUNAKAN ALGORITMA BISECTING K-MEANS

By:

Celvin Gazinda

Technology always developed every year for support human's daily activity. In this digital era, many people have been helped by technological advances, especially in electronic sports (eSport) activities. This makes AEGIS participate to helping Dota 2 amateur players to pursue the game so they can become professional players later. However, the problem that the amount of transactions of AEGIS has decreased. Therefore, study was conducted to discover the factors affect AEGIS customer loyalty using Bisecting K-Means algorithm with python machine learning. Then, perform cluster validation using Silhouette Coefficient method to measure off the quality of cluster is. The last step, Customer Relationship Management (CRM) analysis performed to provide recommendations of discounts, promotions for new customers or another recommendations. The results of the application of the Bisecting K-Means algorithm on AEGIS transaction data have a score of 0.81 in 2 clusters and 0.45 in 3 clusters. Based on the results, 2 clusters is the strong clusters. Meanwhile, 3 clusters are said to be weak clusters because there are still data that are not in the actual cluster. After analyzing, cluster 2 is included in the category of loyal customers, so that the recommendation given is in the form of a discount and member card in order to maintain customer loyalty in registering.

Keywoards: Dota 2, machine learning, python, Silhouette Coefficient, CRM

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

DAFTAR ISI

PENGESAHAN TIM PENGUJI SKRIPSI i
PERNYATAAN KEASLIAN ii
LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI iii
MOTTO iv
KATA PENGANTAR v
ABSTRAK vii
ABSTRACT viii
DAFTAR ISI ix
DAFTAR TABEL xii
DAFTAR GAMBAR xiv
BAB I PENDAHULUAN 1
1.1 Latar belakang 1
1.2 Perumusan Masalah 4
1.3 Batasan Masalah 4
1.4 Tujuan Penelitian 4
1.5 Manfaat Penelitian 4
1.5.1 Akademik 5
1.5.2 AEGIS 5
BAB II TINJAUAN PUSTAKA 6
2.1 Tinjauan Penelitian Terdahulu 6
2.2 Teori-Teori Dasar 9
2.2.1 Data Mining 9
2.2.2 Decision Support Systems (DSS) 11
2.2.3 *Preprocessing* 11
2.2.4 *Customer Relationship Management* (CRM) 12
2.2.5 CRM dengan *clustering* 12
2.2.6 *K-Means* 13
2.2.7 *Bisecting K-Means* 14

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

2.2.8 *Silhouette Coefficient* 15
2.2.9 *Machine Learning* 16
2.2.10 *Python* 16
2.2.11 *Scikit Learn* 17
2.3 Integrasi Keilmuan 17
BAB 3 METODOLOGI PENELITIAN 20
3.1 Desain Penelitian 20
3.1.1 Perumusan Masalah 21
3.1.2 Studi Pustaka 21
3.1.3 Pengambilan Data 21
3.1.4 *Preprocessing* 22
3.1.5 Implementasi Algoritma *Bisecting K-Means* 24
3.1.6 Pengujian Metode 25
3.1.7 Analisis dan Evaluasi Hasil 26
3.2 Waktu Penelitian 27
BAB IV HASIL DAN PEMBAHASAN 28
4.1 Pengambilan Data 28
4.2 *Preprocessing* 29
4.2.1 *Data Selection* 29
4.2.2 *Cleaning Data* 31
4.2.3 *Cleansing Data* 33
4.3 Implementasi Algoritma *Bisecting K-Means* 39
4.3.1 Perhitungan Algoritma *Bisecting K-Means* Pertama 40
4.3.2 Perhitungan Algoritma *Bisecting K-Means* Kedua 43
4.4 Hasil Plot *Bisecting K-Means* 45
4.5 Pengujian Metode 46
4.6 Analisis dan Evaluasi Hasil 47
4.6.1 Analisis Hasil 47
4.6.2 Evaluasi Hasil 55
BAB 5 KESIMPULAN DAN SARAN 56
5.1 Kesimpulan 56

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

5.2 Saran 57
DAFTAR PUSTAKA 58

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

DAFTAR TABEL

Tabel 2.1: Tinjauan penelitian terdahulu 6
Tabel 3.1: Atribut data transaksi AEGIS 21
Tabel 3.2: Atribut yang diperlukan dari data Transaksi AEGIS 23
Tabel 4.1: Atribut pada data AEGIS 28
Tabel 4.2: Data selection data transaksi AEGIS 29
Tabel 4.3: Sampel dataset sesi coaching 30
Tabel 4.4: Hasil seleksi data transaksi AEGIS 31
Tabel 4.5: Sampel data tidak valid 32
Tabel 4.6: Irisan data pada atribut MMR 33
Tabel 4.7: hasil mean atribut MMR 34
Tabel 4.8: Dataset dengan penambahan atribut 35
Tabel 4.9: Data coach AEGIS 36
Tabel 4.10: Data hasil preprocessing atribut role 38
Tabel 4.11: Data hasil penggabungan 39
Tabel 4.12: Sampel data transaksi AEGIS pertama iterasi ke-1 40
Tabel 4.13: Data hasil perhitungan Euclidean Distance pertama iterasi ke-1 41
Tabel 4.14: Data penyesuaian dengan cluster pertama iterasi ke-1 41
Tabel 4.15: Data penyesuaian dengan *cluster* pertama iterasi ke-3 42
Tabel 4.16: Perhitungan SSE cluster 1 42
Tabel 4.17: Perhitungan SSE cluster 2 43
Tabel 4.18: Penentuan centroid secara acak 43
Tabel 4.19: Data hasil Euclidean Distance cluster 1 44
Tabel 4.20: Penyesuaian titik centroid cluster 1 iterasi 1 44
Tabel 4.21: Penyesuaian cluster semua data 45
Tabel 4.22: Data hasil clustering algoritma Bisecting K-Means 46
Tabel 4.23: Kategori Loyalitas Pelanggan 49
Tabel 4.24: Data variabel *role* 49

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Tabel 4.25: Data variabel MMR 50
Tabel 4.26: Data variabel tier..... 51
Tabel 4.27: Data variabel tipe paket 53
Tabel 4.28: Data variabel sesi 53

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

DAFTAR GAMBAR

Gambar 3.1: Diagram alur penelitian 20
Gambar 3.2: Diagram alir Bisecting K-Means (Halim et al., 2017) 24
Gambar 3.3: Diagram alur pengujian system 26
Gambar 4.1: Diagram alur pencarian tipe paket 37
Gambar 4.2: Pseudocode Bisecting K-Means 40
Gambar 4.3: Data hasil plot dari algoritma Bisecting K-means 45
Gambar 4.4: Hasil Silhouette Coefficient Bisecting K-Means 47

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

BAB I
PENDAHULUAN

1.1 Latar belakang

Teknologi memiliki pengaruh besar mengubah konten game yang menawarkan lingkungan imersif dengan lebih signifikan (Macey & Hamari, 2018). Ruang *game* menjadi lebih luas dan tidak lagi terpusat pada *arcade game* yang hanya bisa dimainkan di rumah, melainkan bisa dimainkan dengan mode *multiplayer game*. Banyaknya orang yang bermain *video game* pada jaman sekarang, baik hanya sekedar untuk bersenang-senang maupun berkompetisi telah menciptakan lapangan kerja baru bagi orang-orang yang ingin fokus menjadi atlet *electronic sport* (eSport) (Wiguna & Wawan, 2016).

Permintaan konsumen akan eSport dan pertumbuhan dari kompetisi *video game* yang terorganisir telah mencuri perhatian yang cukup besar dari industri olahraga, acara, dan hiburan. eSport menjadi populer bagi akademisi manajemen olahraga untuk melakukan penelitian, mendidik siswa dan industri jasa (Funk et al., 2018). Esport diketahui sebagai *video game* kompetitif pada kegiatan profesional dengan mengerahkan jaringan pemain, penonton, penggemar, sponsor, penerbit, pengembang, teknisi, komentator, narator, analisis dan aktor-aktor lainnya dalam konteks industri global yang bekerja untuk khalayak umum yang didedikasikan pada tingkat profesional (Macedo & Falcão, 2020).

Perkembangan komunitas eSport di Indonesia memperlihatkan bahwa Industri eSport mendapatkan dukungan dari berbagai pemangku kepentingan baik dari pihak *developer*, *publisher*, pemerintah maupun masyarakat. Salah satunya upaya pemerintah RI hadir sebagai regulator dan fasilitator dengan menerbitkan Peraturan Menteri Nomor 11 Tahun 2016 tentang Klasifikasi Umum Permainan Interaktif Elektronik, dan hadir pada berbagai kompetisi esports, seperti turnamen esport Piala Presiden, Kejuaraan esports Remaja Nasional dan 2019 IEC University Series (Rachman et al., 2020).

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Melihat kejadian yang fenomenal tersebut, maka munculah peluang yang berpotensi untuk dijadikan bisnis dengan mendirikan perusahaan berbentuk Commanditaire Vennootschap (CV) sebagai wadah bagi atlet-atlet eSport di Indonesia (Dewi, 2019). Namun wadah tersebut hanya untuk orang-orang yang memiliki performa pada tingkat semi profesional. Akademi Esport Indonesia (AEGIS) hadir sebagai infrastruktur bagi pemain-pemain amatir yang ingin belajar lebih dalam mengenai *game* yang digeluti dengan para *coach* berpengalaman. *Game* yang tersedia antara lain *Tekken 7, Dota 2 dan Dota Underlord*. Dalam pembelajarannya, AEGIS sudah menyiapkan kurikulum pada setiap game agar pelanggan lebih mudah untuk belajar mengenai ilmu-ilmu yang belum pernah didapatkan selama bermain.

Transaksi pembelian di AEGIS mencapai 339 transaksi dari bulan Januari tahun 2019 sampai bulan Juni tahun 2020. Jumlah transaksi mengalami kenaikan dari bulan Januari hingga mendapat jumlah transaksi tertinggi pada bulan Oktober tahun 2019. Pada bulan November transaksi mengalami penurunan hingga Januari 2020 dengan jumlah transaksi sama seperti pada bulan Januari 2019. Jumlah transaksi tetap konsisten pada bulan Januari sampai Mei tahun 2020. Namun, transaksi mengalami penurunan secara drastis pada bulan Juni 2020. Berdasarkan penjelasan tersebut, AEGIS sedang menghadapi masalah jumlah transaksi yang menurun secara drastic pada tahun 2020. Berdasarkan masalah tersebut dilakukanlah penelitian untuk mencari faktor terkait loyalitas pelanggan yang mempengaruhi kenaikan jumlah transaksi AEGIS menggunakan pendekatan Data Mining. Dengan menerapkan data mining diharapkan hasil dari pengolahan data bisa dilakukan pengambilan keputusan yang tepat oleh AEGIS agar meningkatkan jumlah transaksi sebagai strategi pemasaran selanjutnya.

Data mining digunakan untuk menganalisis dan mencari informasi dari data yang belum teridentifikasi sebelumnya untuk membuat keputusan bisnis yang tepat (Maulida, 2018). *Clustering* merupakan bagian dari metode data mining dan telah diakui sebagai instrumen yang kuat untuk memecahkan kompleksitas permasalahan dari data statistik dan ilmu komputer (Gustientiedina et al., 2019). Clustering digunakan untuk melihat faktor yang mempengaruhi pelanggan tetap loyal melakukan pendaftaran. Harapan dari hasil clustering digunakan sebagai

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

pemecahan masalah yang sedang dihadapi oleh AEGIS sekarang. Terdapat banyak algoritma *clustering* yang digunakan, salah satunya adalah algoritma K-Means.

K-Means menjadi salah satu algoritma yang popular digunakan untuk *clustering*. Popularitasnya disebabkan karena mudah untuk digunakan dan memiliki sistem komputasi yang efisien (Majhi & Biswal, 2018). *K-Means* merupakan algoritma heuristic yang mengelompokan data ke dalam *cluster* K dengan melakukan iterasi untuk mencari hasil *sum squared error* terkecil pada setiap *cluster* (Nishom, 2019). Penelitian ini menggunakan algoritma *Bisecting K-Means* untuk melakukan clustering pada data transaksi AEGIS. *Bisecting K-Means* merupakan metode pengembangan dari algoritma *K-Means* (Abirami & Mayilvahanan, 2016). Algoritma *Bisecting K-Means* menggabungkan teknik hirarki dan metode iteratif *K-Means* untuk mendapatkan hasil *cluster* lebih baik dengan proses komputasi dalam waktu yang lebih sedikit (Swapna & Babu, 2017). Bisecting K-means memiliki keunggulan dari *K-Means* bahwa dengan membagi *cluster* menjadi dua dengan dasar algoritma *K-means* mendapatkan hasil *clustering* yang lebih baik dan lebih konsisten dengan menerapkan *centroid* secara acak (Zhuang et al., 2016).

Algoritma *Bisecting K-Means* digunakan sebagai pengelompokan data untuk mengidentifikasi tingkat loyalitas pelanggan. Setelah melakukan pengelompokan data, selanjutnya mengevaluasi hasil *cluster* dengan metode *Silhouette Coefficient*. Evaluasi hasil *cluster* dilakukan untuk mengukur seberapa baik hasil pengelompokkan data (Azuri et al., 2016). Kemudian dilakukan analisa hasil menggunakan konsep *Customer Relationship Management* (CRM) untuk memberikan rekomendasi yang dibutuhkan oleh AEGIS agar meningkatkan bisnis AEGIS. Karena sekarang adalah era dimana pelanggan menjadi peran yang sangat penting untuk meningkatkan target laba yang diperlukan (Satish & Yusof, 2017). Berdasarkan latar belakang diatas maka dilakukanlah penelitian dengan judul "ANALISIS LOYALITAS PELANGGAN AEGIS PADA PEMBELIAN PAKET COACHING MENGGUNAKAN ALGORITMA BISECTING K-MEANS", harapan dari hasil penelitian ini transaksi AEGIS bisa mengalami peningkatan kedepannya dan menjalin hubungan yang erat dengan pelanggan.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

1.2 Perumusan Masalah

Penjabaran latar belakang diatas dapat dirumuskan masalah menjadi beberapa poin sebagai berikut.

1. Bagaimana mengimplementasikan algoritma Bisecting K-Means dan menghitung performa metode dengan *Silhouette Coefficient* pada data pembelian paket *coaching* AEGIS?
2. Bagaimana menganalisis faktor yang mempengaruhi loyalitas pelanggan pada pembelian paket *coaching* dan rekomendasi CRM yang dibutuhkan oleh AEGIS?

1.3 Batasan Masalah

Penelitian ini mempunyai batasan masalah yang dapat dijelaskan sebagai berikut.

1. Perhitungan algoritma *Bisecting K-Means* menggunakan bahasa pemrograman *python*.
2. Penelitian ini hanya dilakukan pada pelanggan AEGIS.
3. Perhitungan algoritma *Bisecting K-Means* hanya menggunakan data dari pelanggan yang *booking* paket *Dota 2.*
4. Cluster yang digunakan hanya 3 sebagai kategori loyalitas pelanggan yang dijabarkan menjadi pelanggan loyal, moderat dan tidak loyal.

1.4 Tujuan Penelitian

Tujuan dari penelitian ini dijabarkan menjadi beberapa poin sebagai berikut.

1. Mengimplementasikan algoritma Bisecting K-Means dan menghitung performa metode dengan *Silhouette Coefficient* pada data pembelian paket *coaching* AEGIS
2. Menganalisis faktor yang mempengaruhi loyalitas pelanggan pada pembelian paket *coaching* dan rekomendasi CRM yang dibutuhkan oleh AEGIS

1.5 Manfaat Penelitian

Dari penelitian yang dilakukan diharapkan memiliki manfaat sebagai berikut.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

1.5.1 Akademik

1. Pengabdian kepada masyarakat dengan mengimplementasikan ilmu pengetahuan untuk membantu menyelesaikan masalah yang sedang dihadapi.
2. Membantu untuk mengembangkan metode yang diterapkan dari ilmu pengetahuan yang didapat.
3. Sebagai referensi untuk dilakukan pengembangan pada penelitian selanjutnya.

1.5.2 AEGIS

1. Membantu menganalisis permasalahan yang sedang dihadapi oleh AEGIS.
2. Membantu meningkatkan bisnis AEGIS agar jumlah transaksi pembelian dapat meningkat kedepannya.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

BAB II
TINJAUAN PUSTAKA

2.1 Tinjauan Penelitian Terdahulu

Demi mendapatkan pengetahuan dan interpretasi, maka dilakukan tinjauan pada penelitian terdahulu yang memiliki relevansi dengan penelitian yang pada nantinya akan dilakukan. Tinjauan pustaka dari penelitian terdahulu dapat dijabarkan pada Tabel 2.1 sebagai berikut.

Tabel 2.1: Tinjauan penelitian terdahulu

No.	Judul	Hasil Penelitian	Korelasi dengan penelitian yang dilakukan
1.	*A Framework for Outlier Detection Using Improved Bisecting k-Means Clustering Algorithm* (Swapna & Babu, 2017)	Metode *Bisecting K-Means* memberikan hasil *cluster* yang efisien tanpa outlier, dengan waktu komputasi yang lebih sedikit. Metode ini sangat membantu untuk mengembangkan sistem diagnosa hati otomatis berbasis perangkat lunak.	Menggunakan algoritma *Bisecting K-Means* untuk melakukan pengelompokan data.
2.	*A Limited-Iteration Bisecting K-means for Fast Clustering Large*	Dengan tiga iterasi menghasilkan efisiensi komputasi yang lebih tinggi bila dikombinasikan dengan *Bisecting K-Means*.	

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

	Datasets (Zhuang et al., 2016)	Dengan membatasi iterasi, *Bisecting K-means* memiliki kemampuan untuk menghasilkan keakuratan yang lebih tinggi dengan tetap mempertahankan kualitas clustering.	
3.	Pengelompokan Kabupaten/Kota Di Pulau Jawa Berdasarkan Pembangunan Manusia Berbasis Gender Menggunakan Bisecting K-Means (Azuri et al., 2016)	*Cluster* yang dilakukan dibedakan menjadi dua, yaitu *cluster* untuk jenis kelamin laki-laki dan perempuan. Hasil penelitian menunjukkan bahwa nilai silhouette pada kedua variabel dikatakan lemah. Hal ini disebabkan karena masih terdapat Kabupaten/Kota tidak berada pada cluster yang seharusnya.	Menggunakan metode *Silhouette Coefficient* sebagai validasi cluster untuk mengukur seberapa baik pengelompokkan data dari algoritma *Bisecting K-Means*
4.	*Tweet Clustering Using Bisecting K-Means* (Banu & C.R, 2017)	*Bisecting K-Means clustering* bekerja secara efektif untuk mengelompokkan tweet yang sama dan juga menunjukkan bahwa kinerjanya lebih tinggi	

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

		daripada *K-Means clustering*.	
5.	Model Segmentasi Pelanggan Dengan *Kernel K-Means Clustering* berbasis *Customer Relationship Management* (Lubis, 2016)	Dari penelitian tersebut, didapatkan didapatkan acuan sebagai kaidah yang dapat digunakan sebagai penilaian pelanggan terbaik untuk memberi manfaat bagi pihak perusahaan saat mengambil keputusan. *Clustering* digunakan sebagai perbandingan menggunakan dua dan tiga *centroid*. Hasilnya menunjukkan bahwa nilai SSE pada dua nilai *centroid* lebih kecil daripada tiga nilai *centroid*.	Menentukan jumlah cluster berdasarkan tingkat loyalitas pelanggan.
6.	Aplikasi *Data Mining* Dengan Menggunakan Metode *K-Means* Untuk Mengetahui Pola Loyalitas Pelanggan Pada Ali Bakri Konveksi	Dari hasil penelitian pembuatan aplikasi data mining digunakan untuk menganalisa pola loyalitas pelanggan dari data transaksi pelanggan Ali Bakri Konveksi dengan atribut frekuensi pembelian, transaksi terakhir dan total	

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

	(Pratama et al., 2019)	transaksi yang telah dilakukan oleh pelanggan.	
7.	Strategi *Customer Relationship Management* Untuk Meningkatkan Loyalitas pelanggan dan Volume Penjualan Menggunakan Teknik Clustering *K-Means* (Joyendri, 2017)	Hasil kluster memberikan informasi daerah mana saja yang perlu dilakukan tindakan khusus, kemudian dihubungkan dengan program Customer Relationship Management (CRM) seperti pemberian diskon dan memberikan kartu member.	Memberikan rekomendasi CRM sebagai strategi pemasaran agar jumlah transaksi pembelian meningkat kedepannya.
8.	*Cluster Analysis Using Data Mining Approach to Develop CRM Methodology to Assess The Customer Loyalty* (Hosseini et al., 2010)	Hasil dari klustering digunakan untuk menilai kesetiaan pelanggan yang nantinya digunakan untuk strategi pemasaran dibandingkan dengan pemilihan acak yang biasa digunakan oleh sebagian besar perusahaan di Iran.	

2.2 Teori-Teori Dasar

2.2.1 Data Mining

Data mining banyak diterapkan untuk menampilkan nilai-nilai dari data operasi yang dibangun dalam jumlah besar dengan tujuan untuk meningkatkan

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

kinerja sistem operasi yang telah dibuat (Zhao et al., 2019). Data berasal dari kumpulan fakta dan diolah menjadi suatu informasi dengan mining membentuk pola-pola data untuk menyelesaikan suatu masalah.

Data Mining bertujuan untuk membangun sebuah model yang bersifat deskriptif atau prediktif, menampung sejumlah data yang besar, membantu menggali informasi dan Tren yang tersembunyi dalam data (Hemeida et al., 2019).

Beberapa proses dalam *Data Mining* sebagai berikut (Mardi, 2019):

2.2.1.1 *Description*

Mencari pola yang terdapat kecenderungan dalam data kemudian digambarkan agar mudah dipahami. Contohnya petugas KPU mengumpulkan suara tidak bisa menemukan fakta bahwa calon mana yang memiliki pendukung sedikit.

2.2.1.2 *Classification*

Proses mendeteksi sebuah bentuk yang berfungsi dan mengklasifikasikan data pada setiap kelas. Contohnya, penggolongan pangkat dapat kategorikan menjadi tiga, yaitu gaji tinggi, gaji sedang dan gaji rendah.

2.2.1.3 *Estimation*

Estimation hampir mirip dengan *Classification*, kecuali untuk target variabel *Estimation* yang arahnya lebih ke numerik. Contohnya, mahasiswa bisa melihat Indeks Prestasi Kumulatif selama masa perkuliahan.

2.2.1.4 Prediction

Prediction hampir sama dengan *Classification* dan *Estimation*, yaitu mengklasifikasi berdasarkan prediksi nilai di masa datang. Contohnya:

1. Memprediksi harga sembako berupa beras untuk tiga bulan mendatang.
2. Memprediksi jumlah pengangguran pada lima tahun yang akan datang.
3. Memprediksi kenaikan persentase angka kecelakaan lalu lintas pada masa mendatang apabila batas kecepatan minimal dinaikkan.

2.2.1.5 Clustering

Clustering melakukan pengelompokkan record data tanpa memperhatikan kelas-kelas tertentu dan membentuk tipe objek yang memiliki kemiripan pada

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

perekaman data dalam satu kelas yang bernilai maksimal dan kelas lain yang bernilai minimal. Contoh *Clustering*, yaitu:

1. Membuat kelompok kelas-kelas konsumen yang didapatkan dengan tujuan memasarkan suatu produk dari badan usaha yang mempunyai anggaran kecil.
2. Membuat kelompok untuk mengkategorikan baik dan buruk terhadap perilaku finansial.
3. Melakukan *Clustering* pada kemiripan gen terhadap ekspresi dari setiap gen dalam jumlah besar.

2.2.1.6 *Association*

Association bertugas mencari atribut yang terlihat pada waktu yang bersamaan. Contohnya:

1. Meneliti jumlah *customer* yang memberikan tanggapan positif dari penawaran untuk meningkatkan layanan pada perusahaan telekomunikasi seluler.
2. Meneliti barang yang sering dibeli dan tidak pernah dibeli dalam satu waktu oleh *customer*.

2.2.2 Decision Support Systems (DSS)

Decision Support Systems bertugas dalam hal pengambilan keputusan secara fleksibel, interaktif dan nantinya bisa dikembangkan untuk menemukan solusi pada persoalan khusus yang tidak terstruktur, mengaplikasikan data, membuat *User Interface* yang mudah dipahami kemudian menggabungkan pemikiran dalam pengambilan keputusan pada persoalan tersebut (Suryani et al., 2017). DSS meliputi dua hal yang penting, yaitu sistem informasi dan keputusan. Sistem Informasi terdiri dari kumpulan tahapan yang memiliki prosedur formal pada proses pertama data dikelompokkan, kedua data dioperasikan untuk memberikan informasi yang berguna kepada *user.* Keputusan merupakan sistem untuk menemukan solusi dari suatu masalah (Umar et al., 2018).

2.2.3 *Preprocessing*

Preprocessing bertujuan untuk melakukan pembersihan data yang bersih dan meningkatkan akurasi analisis *cluster* dengan menghilangkan noise data dan

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

pengisian pada *missing value* sehingga data menjadi terstruktur agar meningkatkan akurasi pada pengelompokan data (Swapna & Babu, 2017). Beberapa tahapan dari preprocessing dapat dilihat sebagai berikut (Rahman et al., 2017).

1. *Data selection*

   Data selection digunakan untuk melakukan seleksi data yang akan digunakan untuk proses *clustering.*

2. *Cleaning Data*

   Cleaning data bertujuan untuk membersihkan noise dan data yang tidak konsisten pada dataset.

3. *Cleansing Data*

   Data dilakukan *cleansing* untuk melakukan transformasi data menjadi lebih sesuai untuk dilakukan proses *clustering.*

2.2.4 *Customer Relationship Management* (CRM)

Customer Relationship Management (CRM) menjadi irisan bagian instrumen untuk memperkuat ikatan bisnis oleh perusahaan untuk menjalin hubungan yang kuat pada pelanggan (Hardiani et al., 2014). Dengan menggunakan CRM, perusahaan dapat mengetahui keinginan dari pelanggan sehingga membentuk ikatan emosional untuk mewujudkan ikatan bisnis yang kuat dan terbuka, serta menciptakan ikatan yang kuat dari perusahaan pada pelanggan. Apabila bisnis dikelola dengan baik secara alami bisa membentuk hubungan yang baik dengan pelanggan mereka (Saylı et al., 2016). Pada penelitian ini, CRM digunakan untuk memberikan rekomendasi berupa diskon, pencarian pelanggan baru dengan promosi atau rekomendasi lainnya pada hasil kluster yang terbentuk.

2.2.5 CRM dengan *clustering*

Klaster *mining* banyak digunakan untuk CRM dengan menyediakan daftar bisnis yang memungkinkan bisnis untuk menawarkan pelanggan mereka layanan dan produk yang dipersonalisasi (Sharma et al., 2018). Pada lingkungan komersial clustering digunakan, misalnya di bidang berikut:

1. Cross-marketing.
2. Cross-selling.
3. Kustomisasi rencana pemasaran untuk berbagai jenis pelanggan.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

4. Untuk memutuskan metode media apa yang akan digunakan.
5. Identifikasi tujuan belanja

Analisis *Clustering* menjadi jalan sistem pada data mining melalui proses pengiriman objek data kepada kelompok anonim objek tingkat tinggi. Clustering bertugas untuk mensegmentasi populasi yang beragam dalam sejumlah cluster yang lebih homogen. Algoritma pengelompokan diklasifikasikan ke dalam partisi atau hierarkis.

2.2.6 *K-Means*

K-Means adalah teknik pembagian partisi berbasis *prototype* dengan menentukan jumlah klaster berdasarkan *centroid* terdekat (Anitha & Patil, 2020). *K-Means* bertujuan untuk memaksimalkan data yang *similar* pada suatu klaster dan meminimalkan data yang *similar* antar klaster (Asroni, 2015). *K-Means* mengelompokkan n pada pengamatan $X(x_1, x_2, \dots, x_n)$ kedalam k klaster dengan meminimalkan fungsi jarak (Qiao et al., 2019). Perhitungan jarak K-Means menggunakan metode *Euclidean Distance*. *Euclidean Distance* termasuk bagian dari metode kalkulasi pada perhitungan jarak pada beberapa poin *Euclidean Space* (Nishom, 2019). Perhitungan jarak centroid K-Means menggunakan *Euclidean Distance* sebagai proses klasterisasi data. Berikut langkah-langkah dari algoritma K-means:

$$Distance = \sum_{i=1}^{k} \sum_{j=1}^{n} \left\| x_{ij} - C_i \right\|^2 \quad \dots\dots(1)$$

Dimana $C_i = \frac{1}{N_i} \sum_{x \in x_i} x, i = 1,2, \dots, k$ mewakili iterasi pusat klaster.

Berikut adalah algoritma dari *K-Means*:

1. Tentukan klaster *centroid* C_i dengan k sampel acak.
2. Tentukan setiap pengamatan x_i pada pusat klaster terdekat.
3. Hitung kembali dan perbarui setiap klaster $C_i = \frac{1}{N_i} \sum_{x \in x_i} x, i = 1,2, \dots, k$ dimana N_i adalah jumlah elemen pada iterasi klaster.
4. Ulangi tahap dua dan tiga sampai C_i tidak berubah.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

2.2.7 *Bisecting K-Means*

Bisecting K-Means memiliki kombinasi antara *K-Means* dan *hierarchical clustering* yang membagi satu klaster menjadi dua sub kluster pada setiap tahap *bisecting* menggunakan metode *K-Means* sampai *K* kluster tidak berubah (Patil & Khan, 2015). Berikut algoritma dasar *Bisecting K-Means* untuk menemukan *K* kluster (Abirami & Mayilvahanan, 2016):

1. Pilih kluster untuk dibagi.
2. Temukan dua sub kluster menggunakan algoritma dasar *K-Means*.
3. Ulangi tahap dua untuk iterasi ke-n dan ambill hasil pengelompokan yang mempunyai kemiripan.
4. Ulangi tahap satu, dua dan tiga hingga jumlah kluster yang diinginkan tercapai.

Pada penentuan jumlah kluster digunakan perhitungan dengan Sum Square of Error (SSE) dengan rumus sebagai berikut (Syakur et al., 2018).

$$SSE = \sum_{K=1}^{K} \sum_{xi \in S_k} \|X_i - C_k\|_2^2 \quad \ldots\ldots(2)$$

Pada hasil error setiap nilai K akan mengalami penurunan secara perlahan hingga hasil K tersebut stabil dan dilihat pada beberapa nilai K yang mempunyai nilai error paling rendah. Contohnya nilai K=3 nilai errornya menurun drastis pada K=2, K=3 dan K=4. Maka nilai K yang ideal adalah K=3.

Dari langkah-langkah diatas kemudian dapat diimplementasikan pada pseudocode sebagai berikut (Kamat et al., 2017):

```
1.      Tentukan daftar klaster untuk memuat semua data dalam satu kluster
2.      Tentukan klasterNo =1
3.      While(klusterNo !=k)
4.              Bagi dua kluster yang dipilih menggunakan dasar K-Means
                KlusterNo++
                Hitung SSE dari kluster
5.      If (SSE1>SSE2>...>SSEn)
                Tetapkan kluster1 sebagai dataset baru
        Else
                Tetapkan kluster dengan SSE tertinggi sebagai dataset baru
6.      End while
7.      Output K kluster
```


digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

2.2.8 *Silhouette Coefficient*

Metode ini terdiri dari perpaduan antara metode *cohesion* dan *separation* yang diterapkan untuk mengukur kualitas dan kekuatan kluster, yang menentukan seberapa kuat data pada suatu kluster (Anggara et al., 2016). Pengukuran jarak pada dua poin di dalam *cluster* melambangkan metode dari *cohesion*. Sedangkan pada pengukuran jarak pada dua poin diluar *cluster* melambangkan metode dari *separation* (Sholehhudin et al., 2018). Berikut ini ada tiga tahapan dalam menghitung *Silhouette Coefficient* (Wahyuni, I., Auliya, Y. A., & Mahmudy, 2016):

1. Pada tahap pertama, yakni mengukur rata-rata jarak dari data pada satu kluster dengan rumus sebagai berikut.

   $a(i) = \frac{1}{|A|-1} \sum j \in_{A, j \neq i} d(i,j)$ …….....(3)

   Dimana j merupakan data lain yang ada pada satu kluster A dan d(i,j) merupakan jarak antara dokumen i dengan j.

2. Selanjutnya menghitung average jarak dari data ke-i dengan semua poin pada kluster lain yang kemudian diambil nilai paling kecil. Berikut rumusnya.

   $d(i, C) = \frac{1}{|A|} \sum j \in C \; d(i,j)$ ……(4)

   Dimana $d(i, C)$ merupakan jarak rata-rata data ke-i pada semua data kluster lain C, dengan syarat $A \neq C$.

   $b(i) = min \; C \; \neq A \; d(i, C)$ …….(5)

3. Tahap terakhir menghitung nilai *Silhouette Coefficient* pada rumus berikut:

   $s(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))}$ …….(6)

   Rata-rata $s(i)$ merupakan seluruh data pada kluster yang menunjukkan kemiripan dan ketepatan data dalam suatu kluster yang sudah dikelompokkan.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Setelah menghitung *Silhouette Coefficient*, selanjutnya melakukan validasi cluster untuk mengukur kualitas kekuatan cluster dengan standard yang dipaparkan pada tabel 2.2 sebagai berikut (Azuri et al., 2016).

Tabel 2.2 : Interpretasi Nilai *Silhouette*

Silhouette Coefficient Score	Interpretasi
0,71-1,00	*Cluster* yang kuat
0,51-0,70	*Cluster* telah layak atau sesuai
0,26-0,50	*Cluster* yang lemah
≤ 0,25	Tidak dapat dikatakan sebagai *cluster*

2.2.9 *Machine Learning*

Machine Learning merupakan teknik yang membuat komputer belajar dengan mengolah data secara eksplisit dari setiap masalah (Meng et al., 2020). *Machine Learning* telah menjadi kunci dengan kemampuan untuk mengolah kumpulan data besar menjadi informasi yang dibutuhkan (Waring et al., 2020). Selain itu, *Machine Learning* juga menjadi *tools* yang berguna untuk *model, interpret, process* dan *utilize data*, bahkan sampai pada *machine intelligent, data analytics* yang telah banyak digunakan dalam beberapa dekade terakhir dan juga menjadi peran utama untuk mendorong batasan-batasan pada proses *system engineering* (Shang & You, 2019).

2.2.10 *Python*

Python adalah bahasa pemrograman lintas *platform* yang bisa dijalankan pada semua sistem operasi utama (Matthes, 2016). *Python* menjadi pemrograman *open source* populer dan mudah dipahami, serta didukung di berbagai OS. Selain itu, *library* pada *python* memiliki modul khusus diantaranya *threading*, jaringan, database dan sebagainya yang secara umum mudah dipahami (Jaskolka et al., 2019).

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Python banyak digunakan oleh ratusan ribu *developers* di seluruh dunia dalam hal *numeric programming, systems programming, product customization, user interfaces, internet scripting,* dll. Pemrograman *python* masuk pada peringkat 4 atau 5 sebagai bahasa yang populer digunakan oleh publik (Lutz, 2010). Bahkan, pada indeks *PopularitY of Programming Language* (PYPL) bulan Maret tahun 2020 *python* sudah meraih peringkat satu sebagai bahasa pemrograman yang paling populer (Pierre Carbonnelle, 2019).

2.2.11 *Scikit Learn*

Scikit Learn memiliki antarmuka yang mudah dipahami oleh pengguna dan menyediakan banyak implementasi algoritma *machine learning* yang terintegrasi oleh *python* (Pedregosa et al., 2011). Selain itu, scikit learn juga mendukung algoritma machine learning supervised maupun unsupervised learning. Penelitian ini menggunakan *scikit learn* unsupervised learning untuk pengolahan dataset pelanggan AEGIS di python pada nantinya.

2.3 Integrasi Keilmuan

Bisnis yang dilakukan AEGIS merupakan bagian dari perdagangan seperti pada zaman Rasulullah SAW. AEGIS memberikan jasa untuk pembelajaran bagi pemain amatir yang ingin belajar dengan coach yang berpengalaman. Tujuannya agar ilmu yang diberikan oleh *coach* bermanfaat bagi pemain yang nantinya bisa diterapkan untuk menambah kualitas bermain.

Berdasarkan hasil wawancara dengan ustadz Luqman Hakim dari masjid subulussalam mengatakan ada dua firman Allah SWT tentang perdagangan dan mencari ilmu, yaitu surah Al-Baqarah ayat 275 dan surah At-Taha ayat 114. Dalam surah Al-Baqarah ayat 275 dijelaskan sebagai berikut:

اَلَّذِيْنَ يَأْكُلُوْنَ الرِّبٰوا لَا يَقُوْمُوْنَ اِلَّا كَمَا يَقُوْمُ الَّذِيْ يَتَخَبَّطُهُ الشَّيْطٰنُ مِنَ

الْمَسِّۗ ذٰلِكَ بِاَنَّهُمْ قَالُوْٓا اِنَّمَا الْبَيْعُ مِثْلُ الرِّبٰواۘ وَاَحَلَّ اللّٰهُ الْبَيْعَ وَحَرَّمَ الرِّبٰواۗ فَمَنْ

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

جَآءَهٗ مَوْعِظَةٌ مِّنْ رَّبِّهٖ فَانْتَهٰى فَلَهٗ مَا سَلَفَ ۗ وَاَمْرُهٗٓ اِلَى اللّٰهِ ۗ وَمَنْ عَادَ

فَاُولٰٓىِٕكَ اَصْحٰبُ النَّارِ ۚ هُمْ فِيْهَا خٰلِدُوْنَ

Artinya:

"Orang-orang yang memakan riba tidak dapat berdiri melainkan seperti berdirinya orang yang kemasukan setan karena gila. Yang demikian itu karena mereka berkata bahwa jual beli sama dengan riba. Padahal Allah telah menghalalkan jual beli dan mengharamkan riba. Barangsiapa mendapat peringatan dari Tuhannya, lalu dia berhenti, maka apa yang telah diperolehnya dahulu menjadi miliknya dan urusannya (terserah) kepada Allah. Barangsiapa mengulangi, maka mereka itu penghuni neraka, mereka kekal di dalamnya."

Ayat diatas menjelaskan bahwa kegiatan jual beli hukumnya halal dan barang siapa yang menjadikannya riba, maka termasuk golongan penghuni neraka dan kekal di dalamnya. Apabila ia komitmen untuk terus melakukan jual beli dengan cara halal, maka Allah tidak akan menghilangkan pahala orang-orang yang berbuat baik. Lalu pada surat At-Taha ayat 114 dijelaskan sebagai berikut:

فَتَعٰلَى اللّٰهُ الْمَلِكُ الْحَقُّ ۚ وَلَا تَعْجَلْ بِالْقُرْاٰنِ مِنْ قَبْلِ اَنْ يُّقْضٰٓى اِلَيْكَ وَحْيُهٗ ۗ

وَقُلْ رَّبِّ زِدْنِيْ عِلْمًا ۗ

Artinya:

"Maka Mahatinggi Allah, Raja yang sebenar-benarnya. Dan janganlah engkau (Muhammad) tergesa-gesa (membaca) Al-Qur'an sebelum selesai diwahyukan kepadamu, dan katakanlah, "Ya Tuhanku, tambahkanlah ilmu kepadaku." "

Ayat diatas menjelaskan bahwa setiap umat manusia diwajibkan untuk menuntut ilmu. Seperti pada pembelajaran AEGIS, materi kurikulum sudah disediakan untuk orang-orang yang ingin belajar mengenai game yang digeluti dengan harapan ilmu tersebut bisa diterapkan dan untuk perkembangan kedepannya.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

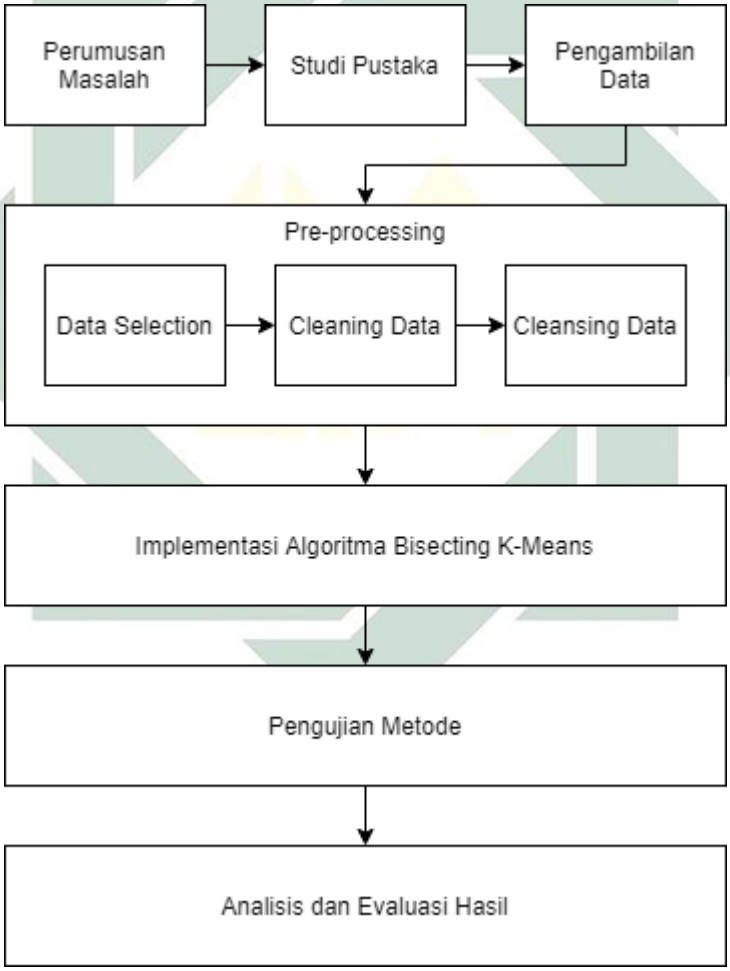
Penelitian ini bertujuan untuk menganalisis loyalitas pelanggan pada pembelian paket *coaching* menggunakan algoritma *Bisecting K-Means* sehingga terlihat faktor apa yang mempengaruhi pelanggan untuk melakukan pembelian paket coaching secara berulang-ulang. Setelah dilakukan analisis maka dilakukanlah evaluasi hasil untuk memberikan solusi cara mempertahankan pelanggan dengan konsep CRM.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

BAB 3
METODOLOGI PENELITIAN

3.1 Desain Penelitian

Alur penelitian dapat digambarkan kedalam kerangka diagram alir yang bertujuan untuk memudahkan dalam pendeskripsian informasi berupa langkah-langkah pada penelitian ini. Berikut ini alur penelitian yang dapat dipaparkan pada Gambar 3.1.



Gambar 3.1: Diagram alur penelitian

Pada Gambar 3.1 diatas dapat dijelaskan diagram alur penelitian dengan langkah-langkah sebagai berikut.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

3.1.1 Perumusan Masalah

Tahap pertama menjelaskan tentang perumusan masalah berdasarkan latar belakang penelitian yang berjudul "ANALISIS LOYALITAS PELANGGAN AEGIS PADA PEMBELIAN PAKET *COACHING* MENGGUNAKAN ALGORITMA *BISECTING K-MEANS"*. Penelitian ini dilakukan karena jumlah transaksi pada Akademi Esport Indonesia (AEGIS) mengalami penurunan. Sehingga, ditemukan rumusan masalah yang mengukur loyalitas pelanggan AEGIS menggunakan metode *Bisecting K-Means.*

3.1.2 Studi Pustaka

Tahap kedua melaksanakan kegiatan studi pustaka berdasarkan kebutuhan dari masalah yang sedang dihadapi oleh AEGIS. Berdasarkan hasil literatur yang dilakukan didapatkan hasil clustering untuk *Customer Relationship Management* (CRM) untuk mengukur loyalitas pelanggan pada AEGIS. Perhitungan clustering yang digunakan adalah *Bisecting K-Means*.

3.1.3 Pengambilan Data

Data yang diperlukan untuk melakukan analisa lebih lanjut didapat dari data Excel transaksi AEGIS dimulai bulan Januari tahun 2019 sampai bulan Juni tahun 2020 dengan atribut pada tabel 3.1 sebagai berikut (Bassam, 2019).

Tabel 3.1: Atribut data transaksi AEGIS

No.	Atribut	Deskripsi
1.	Tanggal	Tanggal transaksi dari *booking* paket *coaching.*
2.	Nama Discord	Nama dari aplikasi Discord *user*. Contoh: CG_LOGIC
3.	*Role*	*Role* artinya peran seseorang dalam permainan dengan deskripsi sebagai berikut. 1. *Mid*, orang yang bertugas bermain di lane mid 2. *Carry*, orang yang bertugas bermain di *lane safelane* 3. *Offlane*, orang yang bertugas bermain di lane *offlane* 4. *Support*, orang yang mempunyai tugas penting untuk membantu tim saat di lane maupun saat war

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

4.	MMR	*MatchMaking Ratio* (MMR) adalah jumlah poin dari tingkat kehebatan seorang pemain pada permainan. Contoh: 3500
5.	*Coach*	Pelatih dari pemain yang sudah *booking* paket *coaching*. Contoh: FlipFlop
6.	Sesi	Jumlah sesi yang diambil. Contoh: 5 sesi
7.	Sesi ke-n	Jumlah sesi ke berapa yang sudah dilakukan dengan format tanggal dan waktu. Contoh: ambil 5 sesi Sesi 1 : 7 (16.00) Sesi 2 : 12 (19.00) Sesi 3 : 12 (21.00) Sesi 4 : 13 (19.00) Sesi 5 : 13 (21.00)

3.1.4 *Preprocessing*

Setelah didapatkan data transaksi dari AEGIS maka dilakukan *preprocessing* data untuk menghindari permasalahan pada saat proses pengolahan data Beberapa tahapan preprocessing yang dilakukan dapat dijabarkan sebagai berikut:

3.1.4.1 *Data Selection*

Tahap pertama melakukan seleksi data dengan menghilangkan data yang tidak dibutuhkan pada proses pengolahan data. Atribut yang tidak diperlukan adalah sesi ke-n. Pada atribut sesi ke-n terdapat data pelanggan yang sudah melakukan sesi *coaching* ke berapa dari sesi yang sudah diambil diawal. Data yang tercatat berupa tanggal dan waktu dari setiap sesi yang sudah dilakukan. Demi menghindari permasalahan saat melakukan pengolahan data, mata atribut tersebut

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

dihilangkan. Data yang diperlukan untuk penelitian ini dengan atribut pada Tabel 3.2 sebagai berikut (Bassam, 2019).

Tabel 3.2: Atribut yang diperlukan dari data Transaksi AEGIS

No.	Atribut	Deskripsi
1.	Tanggal	Tanggal transaksi dari *booking* paket *coaching.*
2.	Nama Discord	Nama dari aplikasi Discord *user*. Contoh: CG_LOGIC
3.	*Role*	*Role* artinya peran seseorang dalam permainan dengan deskripsi sebagai berikut. 5. *Mid*, orang yang bertugas bermain di lane mid 6. *Carry*, orang yang bertugas bermain di *lane safelane* 7. *Offlane*, orang yang bertugas bermain di lane *offlane* 8. *Support*, orang yang mempunyai tugas penting untuk membantu tim saat di lane maupun saat war
4.	MMR	*MatchMaking Ratio* (MMR) adalah jumlah poin dari tingkat kehebatan seorang pemain pada permainan. Contoh: 3500
5.	*Coach*	Pelatih dari pemain yang sudah *booking* paket *coaching*. Contoh: FlipFlop
6.	Sesi	Jumlah sesi yang diambil. Contoh: 5 sesi

3.1.4.2 *Cleaning Data*

Setelah melalui tahap seleksi data, selanjutnya menghilangkan data yang tidak valid pada tabel 3.2. Pada data transaksi AEGIS, ditemukan bahwa ada beberapa data yang dapat mengganggu jalannya proses pengolahan data pada nantinya. Data tersebut diantaranya, data transaksi dengan paket *Team Coaching* dan *Underlord*, serta data transaksi dengan sesi yang tidak valid. Sehingga,

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

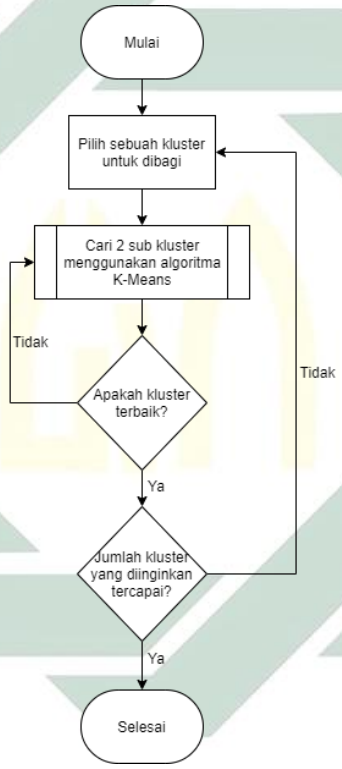
beberapa data tersebut dihilangkan agar tidak mengganggu proses pengolahan data pada saat dijalankan oleh *python* pada nantinya.

3.1.4.3 *Cleansing Data*

Tahap terakhir melakukan *cleansing data* dengan menambahkan, mengubah format, menghapus dan melakukan pengisian pada data yang kosong.

3.1.5 Implementasi Algoritma *Bisecting K-Means*

Algoritma *Bisecting K-Means* dapat dijabarkan melalui diagram alir pada Gambar 3.2 sebagai berikut.



Gambar 3.2: Diagram alir *Bisecting K-Means* (Halim et al., 2017)

Perhitungan *centroid Bisecting K-Means* menggunakan metode *Euclidean Distance* yang berdasar pada algoritma *K-means*. Penjelasan alur dari diagram alir algoritma *Bisecting K-Means* pada gambar 3.2 akan dijelaskan menjadi beberapa langkah sebagai berikut.

1. Pertama, pilih sebuah kluster untuk dilakukan pembagian menjadi dua sub kluster.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

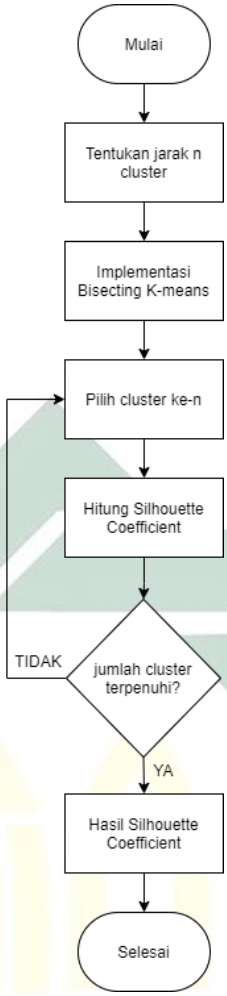
2. Kedua, cari dua sub kluster menggunakan dasar algoritma *K-Means* yang memiliki kemiripan tinggi dan mempunyai SSE terendah.
3. Ketiga, ulangi langkah kedua untuk mencari dua sub kluster yang memiliki kemiripan tinggi dan mempunyai SSE tertinggi.
4. Keempat, Ulangi tahap satu, dua dan tiga hingga jumlah kluster yang diinginkan tercapai.

Penentuan jumlah *cluster* dibagi berdasarkan tingkat loyalitas pelanggan menjadi tiga bagian, diantaranya pelanggan loyal, moderat dan tidak loyal. Hasil dari *clustering* oleh algoritma *Bisecting K-Means* nantinya akan dilakukan pengujian system dan dianalisa untuk melihat faktor apa yang mempengaruhi loyalitas pelanggan AEGIS.

3.1.6 Pengujian Metode

Pengujian sistem dilakukan dengan menghitung *score Silhouette Coefficient* untuk mengukur kualitas dan kekuatan cluster, yang menentukan seberapa baik objek pada setiap kluster. Langkah-langkah pengujian sistem dapat dijabarkan pada gambar 3.3 sebagai berikut.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id



Gambar 3.3: Diagram alur pengujian system

Diagram alur pengujian system pada gambar 3.3 diatas dapat diuraikan sebagai berikut.

1. Menentukan jarak cluster untuk melakukan perhitungan pada setiap n *cluster*.
2. Pengolahan data dengan algoritma *Bisecting K-Means*.
3. Pilih cluster ke-n untuk melakukan perhitungan.
4. Mengitung hasil cluster ke-n dengan *Silhouette Coefficient*.
5. Ulangi langkah 3 dan 4 bila jumlah cluster belum terpenuhi.
6. Menampilkan hasil perhitungan *Silhouette Coefficient* pada setiap cluster.

3.1.7 Analisis dan Evaluasi Hasil

Tahap terakhir melakukan analisis dan evaluasi hasil dari *clustering Bisecting K-Means* untuk melihat faktor yang membuat pelanggan tetap loyal untuk membeli paket *coaching* di AEGIS. Beberapa variabel yang digunakan untuk

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

menganalisis tingkat loyalitas pelanggan dapat dijabarkan pada tabel 3.3 sebagai berikut.

No.	Variabel	Atribut	Deskripsi
1.	*Role*	1. *Mid* 2. *Carry* 3. *Offlane* 4. *Support*	Melakukan penjumlahan pada setaip *role* berdasarkan jumlah pendaftaran yang dilakukan.
2.	MMR	1. Rendah 2. Sedang 3. Tinggi	Jumlah MMR dikategorikan menjadi 3 untuk dilakukan penjumlahan pada setiap tingkat berdasarkan jumlah pendaftaran yang dilakukan.
3.	*Tier*	1-4	Melakukan penjumlahan pada setaip *tier* berdasarkan jumlah pendaftaran yang dilakukan.
4.	Tipe Paket	Reguler dan Premium	Melakukan penjumlahan pada tipe paket regular dan premium berdasarkan jumlah pendaftaran yang dilakukan.
5.	Sesi	1,3,5	Melakukan penjumlahan pada setaip sesi yang diambil berdasarkan jumlah pendaftaran yang dilakukan.

Setelah itu, diberikan rekomendasi dengan *Customer Relationship Management* (CRM) untuk melakukan strategi pemasaran yang tepat agar transaksi pembelian AEGIS semakin meningkat kedepannya dan memperkuat hubungan dengan pelanggan agar pelanggan tetap loyal untuk melakukan pembelian paket *coaching*..

3.2 Waktu Penelitian

Penelitian ini dengan judul "Analisis Loyalitas Pelanggan AEGIS Pada Pembelian Paket *Coaching* Menggunakan Algoritma *Bisecting K-Means*" dilaksanakan pada bulan Januari tahun 2019 sampai bulan Juni tahun 2020. Penelitian ini dilakukan di perusahaan AEGIS, Surabaya.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

BAB IV
HASIL DAN PEMBAHASAN

Pada bab ini dipaparkan hasil dan pembahasan pengolahan data dari metodologi penelitian pada bab sebelumnya. Langkah-langkah yang akan dilakukan yaitu mulai dari persiapan data, pengembangan sistem dengan algoritma *Bisecting K-means* dan pengujian validitas kekuatan cluster dengan metode *Silhouette Coefficient*, serta analisa hasil dari hasil pengolahan untuk melihat faktor apa yang mempengaruhi pelanggan AEGIS ingin melakukan pembelian paket *coaching*. Setelah itu, dilakukan evaluasi hasil untuk pemberian rekomendasi *Customer Relationship Management* (CRM) untuk mempersiapkan strategi pemasaran yang tepat dengan harapan transaksi pembelian AEGIS semakin meningkat kedepannya.

4.1 Pengambilan Data

Bagian awal dilakukan pengambilan data secara manual pada data transaksi AEGIS di Excel. Data diambil dari bulan Januari tahun 2019 sampai dengan bulan Juni tahun 2020 dan memiliki total 339 data. Pada tabel 4.1 berikut akan dijelaskan secara detail macam-macam atribut yang didapatkan dari proses pengambilan data (Bassam, 2019).

Tabel 4.1: Atribut pada data AEGIS

Atribut	Tipe Data	Contoh Data
Tanggal	*Date*	24/01/2019
Nama Discord	*String*	CG_LOGIC
Role	*String*	*Mid, Carry, Offlane, Support*
MMR	*Integer*	3500
Coach	*String*	FlipFlop
Sesi	*Integer*	1,3,5
Sesi ke-n	*Integer*	1,2,3,4,5

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

4.2 *Preprocessing*

Tahap kedua dilakukan *preprocessing* untuk menghilangkan data yang tidak diperlukan, pengisian *missing value* dan mengubah format data yang terdapat pada data transaksi AEGIS agar data dapat diolah oleh python. Beberapa tahapan yang dilakukan sebagai berikut.

4.2.1 *Data Selection*

Bagian ini atribut data transaksi AEGIS yang tidak dibutuhkan dilakukan seleksi data agar proses clustering dari algoritma *Bisecting K-Means* bisa dilakukan dengan mudah kedepannya. Berikut ini penjabaran yang dilakukan pada tabel 4.2 (Bassam, 2019).

Tabel 4.2: *Data selection* data transaksi AEGIS

Atribut	Kondisi	Deskripsi
Tanggal	Tidak ada perubahan	Tidak dibutuhkan pada proses clustering
Nama Discord	Tidak ada perubahan	Tidak dibutuhkan pada proses clustering
Role	Tidak ada perubahan	*Role* artinya peran seseorang dalam permainan dengan deskripsi sebagai berikut. 1. *Mid*, orang yang bertugas bermain di lane *mid* 2. *Carry*, orang yang bertugas bermain di *lane safelane* 3. *Offlane*, orang yang bertugas bermain di *lane offlane* 4. *Support*, orang yang mempunyai tugas penting untuk membantu tim saat di lane maupun saat war
MMR	Tidak ada perubahan	*MatchMaking Ratio* (MMR) adalah jumlah poin dari tingkat kehebatan seorang pemain pada permainan. Contoh: 3500

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Coach	Tidak ada perubahan	Tingkat *coach* berdasarkan kemampuan *coach*. Semakin tinggi tingkatnya maka semakin naik harganya. Berikut adalah penjabaran dari harga *coach* dari masing-masing tier.
Sesi	Tidak ada perubahan	Jumlah sesi yang diambil oleh pelanggan. Jumlah sesi yang tersedia diantaranya, satu, tiga dan lima.
Sesi ke-n	Atribut dihilangkan	Jumlah sesi yang telah dilakukan oleh pelanggan.

Pada tabel diatas atribut sesi ke-n dihilangkan dari dataset karena memiliki isi data yang membuat *python* tidak bisa memproses pengolahan data. Isi data tersebut berupa tanggal dan waktu pelanggan telah melakukan sesi *coaching*. Kesepakatan waktu *coaching* dilakukan atas persetujuan antara dua belah pihak, yaitu antara *coach* dengan pelanggan. Contoh data dari sesi ke-n dapat dilihat pada tabel 4.3 berikut (Bassam, 2019).

Tabel 4.3: Sampel dataset sesi *coaching*

SESI	SESI 1	SESI 2	SESI 3	SESI 4	SESI 5
5	19/07 (13.00)	29/08 (15.30)	21/08 (13.30)	21/08 (16.60)	29/08 (13.30)
1	19/07 (20.30)				
5	23/07 (20.30)	24/07 (19.00)	24/07 (20.30)	30/07 (19.00)	05/08 (20.00)
3	23/07 (13.00)	23/07 (14.30)			

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

5	17/07 (21.00)	31/08 (22.00)	14/10 (19.00)		

Setelah proses seleksi data dilakukan maka didapatkan hasil pada tabel 4.4 sebagai berikut (Bassam, 2019).

Tabel 4.4: Hasil seleksi data transaksi AEGIS

Atribut	Tipe Data	Contoh Data
Tanggal	*Date*	24/01/2019
Nama Discord	*String*	CG_LOGIC
Role	*String*	*Mid, Carry, Offlane, Support*
MMR	*Integer*	3500
Coach	*String*	FlipFlop
Sesi yang diambil	*Integer*	1,3,5

4.2.2 *Cleaning Data*

Pembersihan data dilakukan untuk menghilangkan beberapa data transaksi AEGIS yang tidak valid dan tidak sesuai dengan kriteria atribut. Hal ini dilakukan agar proses *clustering* dapat berjalan dengan lancar dan tidak mengalami masalah karena data yang noisy. Pada proses pencarian data, ditemukan bahwa atribut tipe paket dan sesi mempunyai beberapa data yang bisa menyebabkan noisy. Maka dilakukan proses penghilangan pada beberapa data tersebut. Beberapa data tersebut dapat dilihat pada tabel 4.5 sebagai berikut (Bassam, 2019).

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Tabel 4.5: Sampel data tidak valid

TANGGAL	NAMA DISCORD	ROLE	MMR	COACH	SESI
31/08/2019	Universitas Indonesia	TEAM COACHING		FlipFlop	5
22/09/2019	STIKOM	TEAM COACHING		Hibana	2
08/10/2019	michaeledward8	Support	3400	Hibana	2
11/10/2019	Fadil120996	Carry	1600	Hibana	10
07/11/2019	ITS	TEAM COACHING		FlipFlop	4
03/11/2019	christwifto	Underlord	White Spire	Faza	1

Pada bagian pertama terdapat atribut role dan MMR terdapat data dengan isi *team coaching*. Paket *team coaching* merupakan paket khusus bagi tim Dota 2 yang ingin melakukan sesi *coaching*. Data ini tidak bisa digunakan karena atribut *role* seharusnya berisi kategori antara *mid*, *carry*, *offlane* atau *support*. Sedangkan, data *team coaching* memiliki 5 *role* antara lain *mid*, *carry*, *offlane*, *support* dan *hard support*. Sama halnya dengan atribut MMR pada data *team coaching* terdiri dari 5 orang yang memiliki jumlah MMR berbeda. Sedangkan, atribut MMR hanya memiliki 1 jumlah MMR dari setiap pelanggan sehingga beberapa data tersebut dihilangkan.

Pada bagian kedua terdapat atribut MMR berisi *White Spire* dengan pelanggan bernama christwifto. *White Spire* merupakan tingkat kehebatan seorang pemain sama halnya dengan MMR. Beberapa data dengan isi *team coaching* dan *White Spire* tidak bisa digunakan karena atribut *role* seharusnya berisi kategori seperti pada tabel 4.4 dan MMR berupa angka. Sehingga, beberapa data tersebut dihilangkan dari dataset karena dapat mengganggu berlangsungnya proses pengolahan data oleh *python*.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Kemudian, pada bagian ketiga melakukan pembersihan data pada atribut sesi. Atribut sesi sudah ditentukan memiliki kategori satu, tiga dan lima. Selain daripada itu data dianggap tidak valid dan dilakukan pembersihan data. Pada atribut sesi ditemukan adanya sesi yang tidak sesuai kategori, yaitu dua, empat dan sepuluh sesi yang dilakukan. Sehingga, data tersebut dihilangkan dari dataset. Setelah melalui beberapa tahapan *cleaning data*, jumlah data transaksi AEGIS menjadi 313 data.

4.2.3 *Cleansing Data*

Tahap terakhir pada bagian *preprocessing* yaitu melakukan cleansing data dengan menambahkan, mengubah format, menghapus dan melakukan pengisian pada data yang kosong. Bagian awal melakukan pengisian pada data yang kosong. Pada atribut MMR terdapat banyak data yang kosong. Irisan data pada atribut MMR dapat dilihat pada tabel 4.6 sebagai berikut (Bassam, 2019).

Tabel 4.6: Irisan data pada atribut MMR

No.	MMR
1	Null
2	Null
3	Null
4	Null
5	Null
…	…
310	3200
311	Null
312	3500
313	2700

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Untuk mengisi data yang kosong tersebut maka dilakukan perhitungan nilai rata-rata (mean) dari semua data atribut MMR dengan rumus sebagai berikut (Ramdan et al., 2016).

$$X = \frac{\Sigma X}{N} \quad \text{.....(7)}$$

Keterangan:

$\bar{X}$ = Mean

$\sum X$ = Nilai data

N = Banyak data

Data dari keseluruhan pada atribut MMR dijumlahkan kemudian dibagi dengan jumlah data. Berikut ini contoh perhitungan menggunakan mean:

$$\bar{X} = \frac{3000 + 2600 + 650 + 2200 + 2900 + \cdots + \mathrm{n}}{313}$$

$$\bar{X} = 3079$$

Hasil dari perhitungan rata-rata MMR yang didapatkan diatas adalah 3079. Kemudian, hasil perhitungan dimasukkan kedalam nilai kosong pada atribut MMR yang dapat dilihat pada tabel 4.7 berikut (Bassam, 2019).

Tabel 4.7: hasil mean atribut MMR

No.	MMR
1	3079
2	3079
3	3079
4	3079
5	3079
…	…
310	3200

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

311	3079
312	3500
313	2700

Setelah melakukan pengisian pada data kosong, selanjutnya menambahkan beberapa atribut pada dataset. Atribut yang ditambahkan diantaranya tipe paket dan pendaftaran ke yang dapat dilihat pada tabel 4.6 sebagai berikut (Bassam, 2019).

Tabel 4.8: Dataset dengan penambahan atribut

Atribut	Tipe Data	Contoh Data
Tanggal	*Date*	24/01/2019
Nama Discord	*String*	CG_LOGIC
Role	*String*	*Mid, Carry, Offlane, Support*
MMR	*Integer*	3500
Coach	*String*	FlipFlop
Tipe paket	*String*	REGULER dan PREMIUM
Sesi	*Integer*	1,3,5
Pendaftaran ke	*Integer*	1,2,3,4,5

Atribut tipe paket didapatkan dari pencarian data berdasarkan *role* dan *coach*. Setiap *coach* memiliki *tier* yang berbeda. Semakin tinggi *tier* dari *coach* tersebut maka harga paket yang disediakan juga semakin mahal. Paket reguler terdiri dari *tier* satu, dua dan tiga, lalu paket premium hanya terdiri dari *tier* empat. Data *coach* dapat dilihat pada tabel 4.9 sebagai berikut (Bassam, 2019).

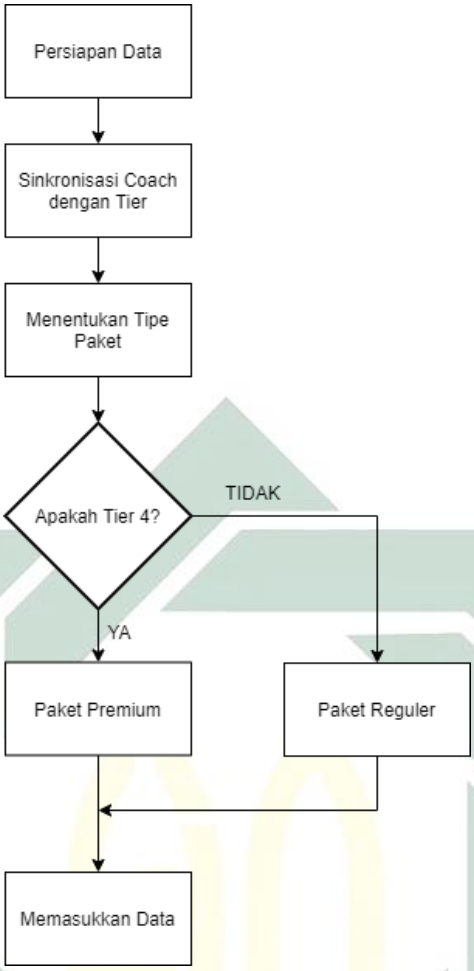
digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Tabel 4.9: Data *coach* AEGIS

No.	Coach	Tier	TIPE PAKET	Role
1.	Panda	4	PREMIUM	All Role
2.	Ilogic	4	PREMIUM	Carry
3.	Varizh	4	PREMIUM	All Role (expert in support)
4.	Kyosh	4	PREMIUM	Mid
5.	Csq	3	REGULER	Mid
6.	Nando	3	REGULER	All Role (expert in core)
7.	alwayssempath	3	REGULER	Support
8.	FlipFlop	3	REGULER	Offlane
9.	Hibana	2	REGULER	Carry
10.	Nedokdokdok	2	REGULER	All Role
11.	raizan	2	REGULER	Carry
12.	badboys	2	REGULER	All Role
13.	William	2	REGULER	Carry/Offlane
14.	Rocketman	2	REGULER	Support
15.	Nidzomus	2	REGULER	All Role
16.	Chadoy	1	REGULER	Support
17.	Shia	1	REGULER	All Role
18.	Divo	1	REGULER	Mid
19.	KaeL	1	REGULER	Offlane
20.	Tortellini	1	REGULER	Support
21.	Veenomon	1	REGULER	Mid/Carry
22.	Agista	1	REGULER	Support/Mid
23.	Destati	1	REGULER	Support/Offlane
24.	EGODEATH666	1	REGULER	Offlane
25.	Wieko_Abimael	1	REGULER	Mid

Tabel diatas digunakan untuk mencari tipe paket yang diambil oleh pelanggan. Langkah-langkah pencarian tipe paket dapat dilihat pada gambar diagram alur 4.1 berikut.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id



Gambar 4.1: Diagram alur pencarian tipe paket

Langkah-langkah diagram alur diatas dapat dijelaskan sebagai berikut:

1. Persiapan data dilakukan sebelum dilakukan pencarian tipe paket.
2. Melakukan sinkronisasi data *coach* dengan *tier* ke berapa kategori dari *coach* tersebut.
3. Setelah dilakukan sinkronisasi data, selanjutnya menentukan tipe paket dari tier *coach*.
4. Jika *coach* memiliki kategori tier keempat, maka termasuk tipe paket premium. Jika *coach* tersebut tidak memiliki kategori keempat, maka termasuk tipe paket reguler.
5. Setelah ditentukan tipe paket, selanjutnya data dimasukkan kedalam atribut tipe paket.

Atribut tipe paket telah terisi, selanjutnya menentukan jumlah pendaftaran yang sudah dilakukan oleh pelanggan. Penentuan jumlah pendaftaran dilakukan

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

pencarian data berdasarkan nama Discord pelanggan. Jumlah pendaftaran diurutkan mulai dari satu sampai jumlah pendaftaran terakhir berdasarkan tanggal transaksi pembelian.

Penambahan data telah selesai, selanjutnya melakukan perubahan format data. Hal pertama yang dilakukan adalah mengubah atribut *coach* menjadi *tier*. Data *coach* memiliki rentang data yang sangat panjang hingga dua puluh lima dan memiliki tipe data *string*. Apabila dilakukan *preprocessing* melalui *scikit learn python* maka data akan memiliki kolom sebanyak dua puluh lima, maka dari itu atribut *coach* tidak efisien untuk digunakan sebagai pengolahan data. Sedangkan, *tier* memiliki rentang data satu sampai dengan empat dan memiliki tipe data *integer*. Pengubahan data dengan cara melakukan sinkronisasi data pada *tier* ke berapa *coach* tersebut, lalu data dimasukkan kedalam atribut *tier*.

Pengubahan atribut *coach* menjadi tier telah dilakukan, selanjutnya mengubah format atribut *role* dari *string* menjadi angka. Atribut *role* terdiri dari empat macam, yaitu *mid, carry, offlane* dan *support*. Agar nantinya data mudah untuk diolah oleh *python* maka dilakukan *preprocessing scikit learn label binarizer*. Hasil preproicessing dapat dilihat pada tabel 4.10 berikut.

Tabel 4.10: Data hasil *preprocessing* atribut role

Index	Carry	Mid	Offlane	Support
0	0	0	0	1
1	0	0	0	1
2	0	1	0	0
3	0	1	0	0
4	0	1	0	0

Data *role* dibuat skala pada setiap anggotanya dengan range angka 0 dan 1 yang dimana angka 0 bila data tidak ada kecocokan data dan angka 1 bila ada kecocokan data. Selanjutnya, data tipe paket yang terdiri dari regular dan premium diubah menjadi angka yang dimana regular menjadi 0 dan premium menjadi 1.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Kemudian data atribut role hasil *preprocessing* digabung dengan data awal. Bagian terakhir melakukan penghapusan atribut yang tidak diperlukan saat proses pengolahan data. Atribut yang dihapus diantaranya tanggal, nama Discord dan *role*. Hasil cleansing data dapat dilihat pada tabel 4.11 sebagai berikut (Bassam, 2019).

Tabel 4.11: Data hasil penggabungan

Variabel	Deskripsi
MMR	*Range* 0 sampai 6000
TIER	1 sampai 4
TIPE PAKET	0 dan 1
SESI	1,3,5
PENDAFTARAN KE	1,2,3,4,5
Mid	0 dan 1
Carry	0 dan 1
Offlane	0 dan 1
Support	0 dan 1

4.3 Implementasi Algoritma *Bisecting K-Means*

Data yang telah dilakukan *preprocessing,* selanjutnya diolah oleh *python* menggunakan algoritma *Bisecting K-Means*. Pada tahap awal menentukan jumlah *cluster* kemudian dibagi berdasarkan tingkat loyalitas pelanggan yang diukur menjadi tiga tingkatan, yaitu pelanggan loyal, moderat dan tidak loyal. Berikut ini *pseudocode* dari algoritma *Bisecting K-Means*.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

```
points = bisecting_kmeans(points, k=3, max_iter=100, verbose=False)
```

Gambar 4.2: *Pseudocode Bisecting K-Means*

4.3.1 Perhitungan Algoritma *Bisecting K-Means* Pertama

Tahap kedua menghitung jarak antara data dengan titik *centroid* pada setiap *cluster* dan menentukan titik *centroid* dengan metode *Euclidean Distance*. Berikut langkah-langkah Algoritma *Bisecting K-Means*.

4.3.1.1 Penentuan *Centroid* Secara Acak Pertama Iterasi Ke-1

Pada iterasi pertama data dianggap pada satu *cluster* kemudian dibagi menjadi dua sub *cluster* dengan menentukan titik *centroid* secara acak. *Centroid* diinisialisasi pada CI (data pertama) dan C2 (data keempat). Sebagai contoh data yang akan digunakan seperti pada tabel 4.12 berikut.

Tabel 4.12: Sampel data transaksi AEGIS pertama iterasi ke-1

No.	MMR	TIER	TIPE PAKET	SESI	PENDAFTARAN KE	Carry	Mid	Offlane	Support	centroid
1.	3000	3	0	1	1	0	1	0	0	C1
2.	2600	1	0	1	1	0	0	0	1	
3.	650	1	0	1	1	0	0	0	1	
4.	2200	1	0	1	1	0	0	1	0	
5.	3900	1	0	1	1	0	0	0	1	
6.	5000	2	0	3	1	0	0	0	1	C2
7.	2600	2	0	1	1	0	0	0	1	
8.	5000	3	0	1	1	0	1	0	0	
9.	3000	3	0	1	1	0	1	0	0	
10.	2000	1	0	1	2	0	0	1	0	

4.3.1.2 Perhitungan *Euclidean Distance* Pertama Iterasi Ke-1

Data pada tabel 4.12 dilakukan perhitungan menggunakan metode *Euclidean Distance* dengan rumus sebagai berikut (Qiao et al., 2019).

$$Distance = \sum_{i=1}^{k} \sum_{j=1}^{n} \left\| x_{ij} - C_i \right\|^2 \quad \text{.....(8)}$$

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Setelah menghitung *Euclidean Distance* maka selanjutnya mengelompokkan data pada masing-masing *centroid* terdekat dengan perhitungan jika Distance C1<Distance C2 maka termasuk *cluster* 1, jika tidak maka termasuk *cluster* 2. Hasil perhitungan Euclidean Distance dapat dilihat pada tabel 4.4 berikut.

Tabel 4.13: Data hasil perhitungan *Euclidean Distance* pertama iterasi ke-1

No.	Distance C1	Distance C2	Cluster
1	0	2000	1
2	400	2400	1
3	2350	4350	1
4	800	2800	1
5	900	1100	1
6	2000	0	2
7	400	2400	1
8	2000	2	2
9	0	2000	1
10	1000	3000	1

4.3.1.3 Pengelompokan Data pertama Iterasi Ke-1

Pada tabel 4.13 masing-masing data terdapat kelompok cluster hasil dari Euclidean Distance. Selanjutnya mengelompokkan data sesuai dengan *cluster* hasil dari *Euclidean Distance* yang dapat dilihat pada tabel 4.14 berikut.

Tabel 4.14: Data penyesuaian dengan *cluster* pertama iterasi ke-1

No.	MMR	TIER	TIPE PAKET	SESI	PENDAFTARAN KE	Carry	Mid	Offlane	Support	Cluster
1.	3000	3	0	1	1	0	1	0	0	1
2.	2600	1	0	1	1	0	0	0	1	1
3.	650	1	0	1	1	0	0	0	1	1
4.	2200	1	0	1	1	0	0	1	0	1
5.	3900	1	0	1	1	0	0	0	1	1
6.	5000	2	0	3	1	0	0	0	1	2
7.	2600	2	0	1	1	0	0	0	1	1
8.	5000	3	0	1	1	0	1	0	0	2
9.	3000	3	0	1	1	0	1	0	0	1
10.	2000	1	0	1	2	0	0	1	0	1

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

4.3.1.4 Pengelompokan Data Pertama Iterasi Ke-2 dan Ke-3

Selanjutnya, melakukan iterasi kedua dan seterusnya dengan menentukan titik centroid seperti yang dilakukan pada iterasi pertama. Pada perhitungan iterasi ketiga titik centroid tidak berubah. Selanjutnya mengelompokkan data sesuai hasil iterasi ketiga yang dapat dilihat pada tabel 4.15 berikut.

Tabel 4.15: Data penyesuaian dengan *cluster* pertama iterasi ke-3

No.	MMR	TIER	TIPE PAKET	SESI	PENDAFTARAN KE	Carry	Mid	Offlane	Support	Cluster
1.	3000	3	0	1	1	0	1	0	0	1
2.	2600	1	0	1	1	0	0	0	1	1
3.	650	1	0	1	1	0	0	0	1	2
4.	2200	1	0	1	1	0	0	1	0	1
5.	3900	1	0	1	1	0	0	0	1	1
6.	5000	2	0	3	1	0	0	0	1	1
7.	2600	2	0	1	1	0	0	0	1	1
8.	5000	3	0	1	1	0	1	0	0	1
9.	3000	3	0	1	1	0	1	0	0	1
10.	2000	1	0	1	2	0	0	1	0	1

4.3.1.5 Perhitungan *Sum Squared Error* (SSE) Pertama

Setelah memasukkan cluster pada setiap data selanjutnya menghitung *Sum Squared Error* (SSE) untuk mencari *cluster* yang memiliki jumlah SSE tertinggi. Berikut ini tabel 4.16 dan 4.17 perhitungan SSE pada setiap cluster.

Tabel 4.16: Perhitungan SSE *cluster* 1

No.	MMR	TIER	TIPE PAKET	SESI	PENDAFTARAN KE	Carry	Mid	Offlane	Support
1.	3000	3	0	1	1	0	1	0	0
2.	2600	1	0	1	1	0	0	0	1
4.	2200	1	0	1	1	0	0	1	0
5.	3900	1	0	1	1	0	0	0	1
6.	5000	2	0	3	1	0	0	0	1
7.	2600	2	0	1	1	0	0	0	1
8.	5000	3	0	1	1	0	1	0	0
9.	3000	3	0	1	1	0	1	0	0
10.	2000	1	0	1	2	0	0	1	0
Rata-rata	3255.56	1.89	0	1.22	1.11	0	0.33	0.22	0.44
SSE	10192222.22	6.89	0	4	1	0	2	1.56	2.22

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

	10182239.33

Tabel 4.17: Perhitungan SSE *cluster* 2

No.	MMR	TIER	TIPE PAKET	SESI	PENDAFTARAN KE	Carry	Mid	Offlane	Support
3.	650	1	0	1	1	0	0	0	1
Rata-rata	0	0	0	0	0	0	0	0	0
SSE	0	0	0	0	0	0	0	0	0
	0								

Pada tabel diatas *Cluster* 1 memiliki nilai SSE lebih tinggi dibandingkan dengan *cluster* 2. Sehingga, *cluster* 1 dilakukan pembagian menjadi dua sub cluster. Pada tahap ini *cluster* sudah terbagi menjadi 3 dan *cluster* 2 dianggap menjadi *cluster* 3.

4.3.2 Perhitungan Algoritma *Bisecting K-Means* Kedua

4.3.2.1 Penentuan *Centroid* Secara Acak Kedua Iterasi Ke-1

Hal pertama sama dilakukan pada proses sebelumnya yaitu, menentukan titik *centroid* secara acak pada *cluster* 1 menjadi dua sub *cluster*. *Centroid* diinisialisasi pada CI (data pertama) dan C2 (data kelima). Sebagai contoh data yang akan digunakan seperti pada tabel 4.18 berikut.

Tabel 4.18: Penentuan *centroid* secara acak

No.	MMR	TIER	TIPE PAKET	SESI	PENDAFTARAN KE	Carry	Mid	Offlane	Support	Centroid
1.	3000	3	0	1	1	0	1	0	0	C1
2.	2600	1	0	1	1	0	0	0	1	
4.	2200	1	0	1	1	0	0	1	0	
5.	3900	1	0	1	1	0	0	0	1	
6.	5000	2	0	3	1	0	0	0	1	C2
7.	2600	2	0	1	1	0	0	0	1	
8.	5000	3	0	1	1	0	1	0	0	
9.	3000	3	0	1	1	0	1	0	0	
10.	2000	1	0	1	2	0	0	1	0	

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

4.3.2.2 Perhitungan *Euclidean Distance* Kedua Iterasi Ke-1

Setelah ditentukan *centroid*, selanjutnya menghitung Euclidean Distance dan menentukan *cluster* pada setiap anggota. Hasil Euclidean Distance dapat dilihat pada tabel 4.19 sebagai berikut.

Tabel 4.19: Data hasil Euclidean Distance *cluster* 1

No.	Distance C1	Distance C2	Cluster
1	0	2000	1
2	400.01	2400	1
4	800	2800	1
5	900	1100	1
6	2000	0	2
7	400	2400	1
8	2000	2.65	2
9	0	2000	1
10	1000	3000	1

4.3.2.3 Pengelompokan Data Kedua Iterasi Ke-1

Selanjutnya mengelompokkan data sesuai dengan cluster hasil Euclidean distance diatas pada tabel 4.20 sebagai berikut.

Tabel 4.20: Penyesuaian titik centroid cluster 1 iterasi 1

No.	MMR	TIER	TIPE PAKET	SESI	PENDAFTARAN KE	Carry	Mid	Offlane	Support	Cluster
1.	3000	3	0	1	1	0	1	0	0	1
2.	2600	1	0	1	1	0	0	0	1	1
4.	2200	1	0	1	1	0	0	1	0	1
5.	3900	1	0	1	1	0	0	0	1	1
6.	5000	2	0	3	1	0	0	0	1	2
7.	2600	2	0	1	1	0	0	0	1	1
8.	5000	3	0	1	1	0	1	0	0	2
9.	3000	3	0	1	1	0	1	0	0	1
10.	2000	1	0	1	2	0	0	1	0	1

Data sudah sesuai dengan cluster masing-masing. Selanjutnya, melakukan iterasi kedua dan seterusnya seperti yang dilakukan pada iterasi pertama. Pada kasus ini, iterasi berhenti pada iterasi ketiga, karena titik centroid tidak berubah.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

4.3.2.4 Pengelompokan Data dari 3 *Cluster*

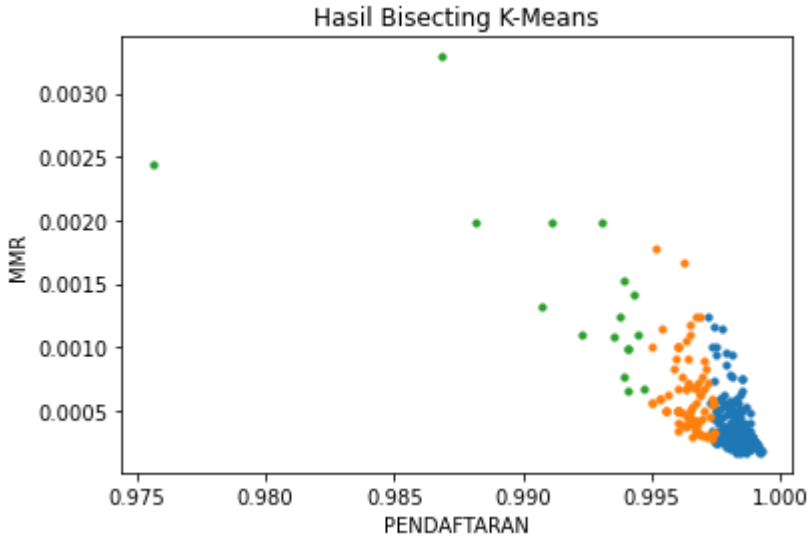
Cluster telah terbagi menjadi 3, selanjutnya mengelompokan data sesuai dengan *cluster*. Berikut ini pengelompokan data berdasarkan cluster pada tabel 4.21.

Tabel 4.21: Penyesuaian *cluster* semua data

No.	MMR	TIER	TIPE PAKET	SESI	PENDAFTARAN KE	Carry	Mid	Offlane	Support	Cluster
1.	3000	3	0	1	1	0	1	0	0	1
2.	2600	1	0	1	1	0	0	0	1	1
3.	650	1	0	1	1	0	0	0	1	3
4.	2200	1	0	1	1	0	0	1	0	1
5.	3900	1	0	1	1	0	0	0	1	2
6.	5000	2	0	3	1	0	0	0	1	2
7.	2600	2	0	1	1	0	0	0	1	1
8.	5000	3	0	1	1	0	1	0	0	2
9.	3000	3	0	1	1	0	1	0	0	1
10.	2000	1	0	1	2	0	0	1	0	1

4.4 Hasil Plot *Bisecting K-Means*

Langkah selanjutnya menampilkan plot hasil dari pengolahan data menggunakan algoritma *Bisecting K-Means* yang dapat dilihat pada gambar 4.3 berikut (Bassam, 2019).



Gambar 4.3: Data hasil plot dari algoritma *Bisecting K-means*

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Dari hasil algoritma *Bisecting K-Means* pada gambar 4.3 diatas, didapatkan hasil pada setiap *cluster* yang beranggotakan 191 pada cluster 1, 104 pada *cluster* 2 dan 16 pada cluster 3 dengan setiap *tier* pada *cluster* masing-masing. Hasil *cluster* dapat dijabarkan pada tabel 4.22 berikut.

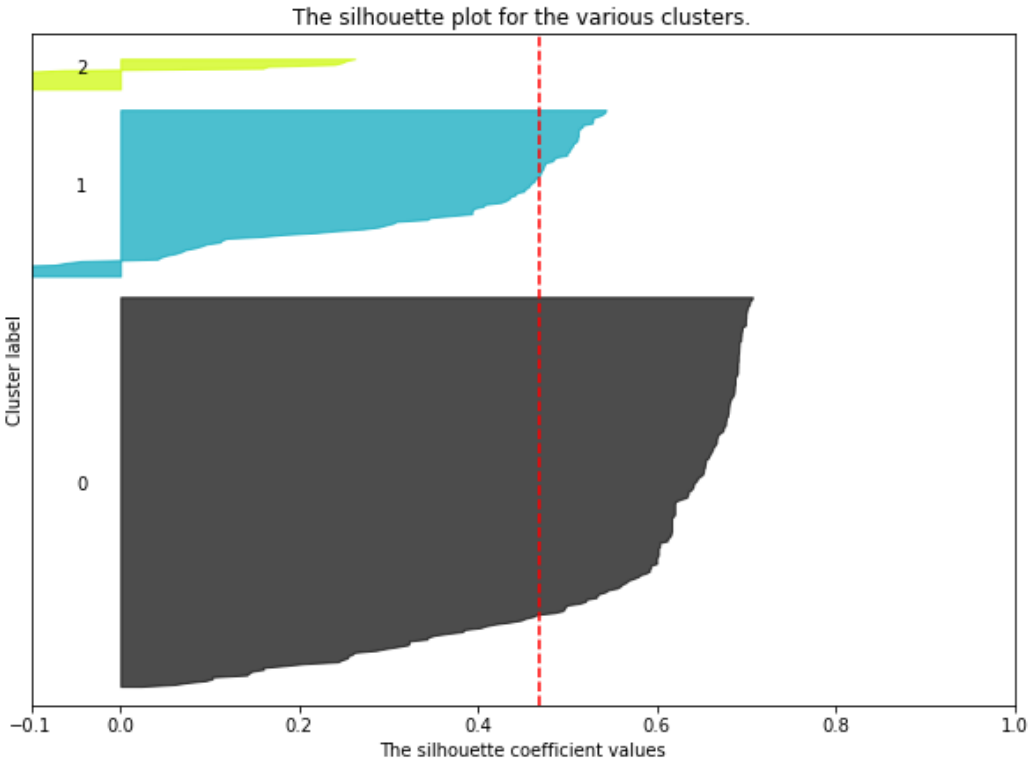
Tabel 4.22: Data hasil clustering algoritma *Bisecting K-Means*

PENDAFTARAN KE	CLUSTER 1	CLUSTER 2	CLUSTER 3
1	158	73	14
2	23	19	2
3	6	7	
4	4	5	
JUMLAH ANGGOTA	191	104	16

4.5 Pengujian Metode

Pada bagian ini hasil yang terbentuk anggota pada setiap *cluster*, selanjutnya dilakukan validasi *cluster* menggunakan *Silhouette Coefficient* untuk mengukur seberapa kuat pengelompokan pada *cluster* tersebut. Dari hasil *Silhouette Coefficient* didapatkan *score* pada 2 *cluster* sebesar 0.81 dan *score* pada 3 cluster sebesar 0,45. Berdasarkan hasil *score* pada setiap *cluster*, 2 *cluster* memiliki *score* paling besar dan dapat dikatakan sebagai *cluster* kuat karena data sudah berada pada *cluster* yang benar. Selain itu, 3 *cluster* juga dikatakan sebagai *cluster* lemah. Hal tersebut disebabkan karena masih banyak data pada *cluster* tersebut yang seharusnya tidak berada pada *cluster* sebenarnya. Hasil plot *Silhouette Coefficient* dapat dilihat pada gambar 4.4 sebagai berikut.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id



Gambar 4.4: Hasil *Silhouette Coefficient Bisecting K-Means*

4.6 Analisis dan Evaluasi Hasil

Bagian terakhir melakukan analisis dan evaluasi dari hasil pengolahan data yang telah dilakukan sebelumnya. Tahap pertama yang dilakukan adalah melakukan analisis terkait loyalitas dari hasil *cluster* pada tabel 4.22 diatas.

4.6.1 Analisis Hasil

Pada bagian ini hasil dari pengelompokan algoritma *Bisecting K-Means* dilakukan analisa dengan menghitung persentase kembalinya pelanggan untuk menentukan kategori loyalitas pelanggan dari setiap *cluster*. Setelah dilakukan perhitungan, kemudian dilakukan pemaparan data dari setiap *cluster* dengan menjabarkan label dari setiap variabel *cluster* untuk mencari faktor terkait loyalitas pelanggan terhadap perusahaan AEGIS.

4.6.1.1 Perhitungan Persentase Loyalitas Pelanggan

Loyalitas pelanggan menjadi peran penting sebagai konsumen untuk tetap bertransaksi pada produk perusahaan secara berulang dan memiliki jangka waktu panjang (Safitri et al., 2019). Untuk mengukur tingkat loyalitas pelanggan maka

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

dilakukan perhitungan persentase kembalinya pelanggan. Variabel loyalitas dibagi menjadi tiga, yaitu loyal, moderat dan tidak loyal. Cluster 3 memiliki jumlah anggota sebanyak 16 dan tidak ada pelanggan yang melakukan pendaftaran diatas 2 kali. Sehingga, *cluster* 3 dimasukkan sebagai kategori pelanggan yang tidak loyal. Selanjutnya, menganalisa *cluster* 1 dan 2 dengan menghitung persentase kembalinya pelanggan pada transaksi dengan rumus sebagai berikut.

Perhitungan *cluster* 1:

$$\frac{\left(\frac{23}{158}\right)+\left(\frac{6}{23}\right)+\left(\frac{4}{6}\right)}{3}$$

$$=\frac{0{,}15+0{,}26+0{,}67}{3}$$

$$=0{,}36$$

Perhitungan *cluster* 2:

$$\frac{\left(\frac{19}{73}\right)+\left(\frac{7}{19}\right)+\left(\frac{5}{7}\right)}{3}$$

$$=\frac{0{,}26+0{,}37+0{,}71}{3}$$

$$=0{,}45$$

Berdasarkan hasil perhitungan diatas *cluster* 2 memiliki nilai lebih besar dengan *score* 0,45 daripada *cluster* 1 dengan score 0,36. Sehingga, *cluster* 2 dimasukkan sebagai kategori pelanggan loyal karena memiliki persentase lebih tinggi kembalinya pelanggan melakukan transaksi dan *cluster* 1 dimasukkan sebagai kategori pelanggan moderat. Hasil analisis dapat dilihat pada tabel 4.23 berikut.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Tabel 4.23: Kategori Loyalitas Pelanggan

Cluster	Kategori
1	Moderat
2	Loyal
3	Tidak Loyal

4.6.1.2 Pemaparan Data Setiap *Cluster*

Tahap kedua dilakukan pemaparan data pada setiap *cluster* hasil dari pengelompokkan algoritma *Bisecting K-Means*. Pemaparan hasil pengelompokan bertujuan untuk mencari terkait faktor yang mempengaruhi loyalitas pelanggan dari variabel yang digunakan antara lain *role*, MMR, *tier*, tipe paket dan sesi. Pertama, variabel *role* pada setiap *cluster* dapat dijabarkan pada tabel 4.24 sebagai berikut (Bassam, 2019).

Tabel 4.24: Data variabel *role*

Cluster	ROLE	PENDAFTARAN KE				Jumlah
		1	2	3	4	
1	Mid	52	6	2	1	61
	Carry	42	8		1	51
	Offlane	15	4	1	1	21
	Support	49	5	3	1	58
2	Mid	15	3	2	2	22
	Carry	21	8	3		32
	Offlane	13	3	1	1	18

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

	Support	24	5	1	2	32
3	Mid	3	1			4
	Carry	6				6
	Offlane	1	1			2
	Support	4				4

Tabel 4.24 diatas pada *cluster* 1 masing-masing *role* pelanggan melakukan pendaftaran hingga 4 kali, namun hanya *role carry* yang transaksinya kosong pada pendaftaran ketiga. Transaksi paling banyak dilakukan oleh pelanggan pada *role mid* dengan jumlah 61 transaksi. Kemudian, *role carry* dan *support* pada *cluster* 2 memiliki jumlah yang sama sebanyak 32 transaksi. Namun, pada *role carry* pelanggan hanya melakukan pendaftaran sampai 3 kali. Lalu, jumlah transaksi pada *cluster* 3 paling banyak pada pelanggan dengan *role carry* sebanyak 6 transaksi. Pendaftaran yang dilakukan hanya sampai 2 kali pada *role mid* dan *offlane* sebanyak 1 transaksi. Dari hasil pemaparan tabel diatas belum ditemukan perbedaan yang signifikan. Selanjutnya, penjabaran data pada variabel MMR dapat dilihat pada table 4.25 sebagai berikut.

Tabel 4.25: Data variabel MMR

Cluster	MMR	PENDAFTARAN KE				Jumlah
		1	2	3	4	
1	0-2000	12				12
	2001-4000	108	14	4	1	127
	4001-6000	38	9	2	3	52
2	0-2000	28	5			33
	2001-4000	45	14	7	5	71

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

	4001-6000					0
3	0-2000	14	2			16
	2001-4000					0
	4001-6000					0

Pada tabel 4.25 jumlah transaksi pada *cluster* 1 paling banyak dilakukan oleh pelanggan dengan MMR antara 2001 sampai 4000 sebanyak 127 transaksi. Diantara *range* MMR diatas hanya pelanggan dengan MMR 0-2000 yang melakukan pendaftaran sekali. Selain itu, MMR 2001-6000 pelanggan melakukan pendaftaran hingga 4 kali. Selanjutnya, pada *cluster* 2 jumlah transaksi paling banyak dilakukan oleh pelanggan dengan MMR antara 2001 sampai 4000 sebanyak 71 transaksi. Lalu, pada MMR antara 0 sampai 2000, pelanggan melakukan transaksi sebanyak 33. Sedangkan, pada MMR 4001-6001 tidak ada pelanggan yang melakukan transaksi. Pendaftaran hingga 4 kali hanya dilakukan oleh pelanggan dengan MMR antara 2001-4000. Terakhir pada *cluster* 3 transaksi hanya dilakukan pada pelanggan yang memiliki MMR antara 0 sampai 2000 sebanyak 16 transaksi. Pelanggan juga melakukan pendaftaran hingga 2 kali sebanyak 2 transaksi.

Berdasarkan hasil penjabaran tabel diatas terdapat perbedaan yang signifikan pada setiap *cluster*. Pelanggan pada *cluster* 1 memiliki kriteria MMR antara 0-6000, *cluster* 2 memiliki kriteria MMR antara 0-4000 dan *cluster* 3 memiliki kriteria MMR antara 0-2000. Selanjutnya, penjabaran data variabel tier dapat dilihat pada tabel 4.26 sebagai berikut

Tabel 4.26: Data variabel *tier*

Cluster	TIER	PENDAFTARAN KE				Jumlah
		1	2	3	4	
1	1	74	6	4		84

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

	2	32	6	1	1	40
	3	49	8	1	3	61
	4	3	3			6
2	1	26	7		1	34
	2	21	7	1		29
	3	21	4	4	4	33
	4	5	1	2		8
3	1	7				7
	2	2				2
	3	3	2			5
	4					0

Tabel 4.26 diatas *cluster* 1 *tier* 1 memiliki jumlah total transaksi lebih banyak daripada *tier* lainnya yaitu sebanyak 84 transaksi. Pelanggan yang melakukan pendaftaran hingga 4 kali hanya pada *tier* 2 dan 3. Tier 4 memiliki jumlah transaksi paling sedikit hanya sebanyak 6 transaksi. Lalu, *cluster* 2 *tier* 1 memiliki jumlah total transaksi lebih banyak dari *tier* lain sebanyak 34 transaksi. Kemudian, pada *tier* 3 memiliki jumlah transaksi hampir sama dengan *tier* 1 sebanyak 33 transaksi. Berdasarkan pendaftaran yang dilakukan pelanggan, jumlah kembalinya pelanggan pada *tier* 3 lebih stabil daripada tier 1. Pada *tier* 3 pelanggan melakukan melakukan pendaftaran sebanyak 4 kali pada pendaftaran 2, 3 dan 4. Sedangkan, pada *tier* 1 pelanggan paling banyak pada pendaftaran kedua dengan jumlah 7 transaksi. Terakhir pada cluster 3 jumlah transaksi paling banyak pada *tier* 1 sebanyak 7 transaksi. Pada tier 4 tidak ada transaksi yang dilakukan oleh pelanggan. Pendaftaran dilakukan hingga 2 kali pada *tier* 3 sebanyak 2 transaksi. Pada penjabaran tabel diatas tidak ditemukan perbedaan yang signifikan pada setiap

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

cluster. Kemudian, pemaparan dilanjutkan pada data variabel tipe paket pada tabel 4.27 sebagai berikut.

Tabel 4.27: Data variabel tipe paket

Cluster	TIPE PAKET	PENDAFTARAN KE				Jumlah
		1	2	3	4	
1	0	155	20	6	4	185
	1	3	3			6
2	0	68	18	5	5	96
	1	5	1	2		8
3	0	14	2			16
	1					0

Berdasarkan tabel 4.27 pelanggan pada *cluster* 1 lebih banyak melakukan pendaftaran pada tipe paket reguler daripada tipe paket premium. Jumlah transaksi pada tipe paket reguler mencapai 185 transaksi sedangkan pada tipe paket premium hanya mencapai 6 transaksi. Kemudian, jumlah transaksi *cluster* 2 tipe paket reguler lebih banyak dari tipe paket premium yaitu sebanyak 96 transaksi. Sedangkan, jumlah tipe paket premium sebanyak 8 transaksi. Lalu, pada cluster 3 transaksi hanya terjadi pada tipe paket reguler sebanyak 16 transaksi. Pendaftaran juga dilakukan sebanyak 2 kali sebanyak 2 transaksi. Dari hasil pemaparan tabel diatas, rata-rata pelanggan mengambil tipe paket reguler. Selanjutnya, memaparkan data variabel sesi pada tabel 4.28 sebagai berikut.

Tabel 4.28: Data variabel sesi

Cluster	SESI	PENDAFTARAN KE				Jumlah
		1	2	3	4	

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

1	1	124	19	6	3	152
	3	18	2		1	21
	5	16	2			18
2	1	14	3	4	2	23
	3	13	4	2	1	20
	5	46	12	1	2	61
3	1	5	1			6
	3	4	1			5
	5	5				6

Pada tabel 4.28 *cluster* 1 paling banyak melakukan pendaftaran pada 1 sesi dengan jumlah 152 transaksi. Sedangkan, pada 5 sesi memiliki jumlah transaksi paling sedikit sebanyak 18 transaksi. *Cluster* 2 jumlah transaksi terbanyak pada 5 sesi sebesar 61 transaksi. Namun, jumlah kembalinya pelanggan melakukan transaksi pada 1 sesi lebih konsisten daripada sesi yang lain. *Cluster* 3 transaksi paling banyak pada 1 sesi dengan jumlah 6 transaksi. Jumlah transaksi yang sama terjadi pada 3 dan 5 sesi sebanyak 5 transaksi. Pendaftaran dilakukan sebanyak 2 kali dengan jumlah 1 transaksi pada 1 dan 3 sesi.

4.6.1.3 Hasil Analisis Faktor Loyalitas Pelanggan

Berdasarkan penjabaran tabel 4.24 sampai dengan 4.28, perbandingan antara *cluster* 1, 2 dan 3 terlihat pada MMR dan jumlah sesi. Pada *cluster* 1 rata-rata pelanggan memiliki MMR antara 0 sampai 6000 dan lebih banyak melakukan transaksi pada 1 sesi. Kemudian pada *cluster* 2, pelanggan rata-rata memiliki MMR antara 0 sampai 4000 dan transaksi paling banyak pada 5 sesi. Lalu, *cluster* 3 rata-rata pelanggan MMR antara 0-2000 dan jumlah transaksi dari semua sesi hapir sama.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

4.6.2 Evaluasi Hasil

Pada bagian ini dilakukan evaluasi dari hasil analisis yang telah dilakukan dengan memberikan rekomendasi *Customer Relationship Management* (CRM). Dengan menggunakan strategi CRM, diharapkan semua transaksi yang melibatkan penyediaan jasa, pemasaran dan penjualan produk kepada pelanggan dapat dilakukan secara terorganisir dan sistematis. Berdasarkan penjabaran pada tabel 4.24 sampai 4.28, strategi CRM yang bisa dilakukan untuk mempertahankan pelanggan tetap melakukan pembelian paket *coaching* bisa dijabarkan sebagai berikut.

4.6.2.1 Strategi CRM pada *Cluster* 1

Strategi CRM yang bisa dilakukan pada *cluster* 1 adalah dengan memberi diskon harga paket (Joyendri, 2017). Hal tersebut bisa dilakukan karena pada *cluster* 1, pelanggan paling banyak membeli paket dengan 1 sesi. Sehingga strategi diskon paket bisa diterapkan pada pembelian paket dengan 3 dan 5 sesi untuk menambah jumlah transaksi AEGIS.

4.6.2.2 Strategi CRM *Cluster* 2

Strategi CRM yang bisa dilakukan pada *cluster* 2 adalah dengan memberi kartu member pada pelanggan loyal agar tetap melakukan transaksi pada perusahaan AEGIS (Joyendri, 2017). Hal tersebut bisa diterapkan karena pelanggan yang loyal memiliki peluang untuk memberikan respon pada produk baru yang diberikan oleh perusahaan. Dengan memberi kartu member pelanggan bisa mendapatkan benefit yang tidak bisa didapatkan pada member biasa.

4.6.2.3 Strategi CRM *Cluster* 3

Strategi CRM yang bisa dilakukan pada *cluster* 3 adalah dengan memberikan promosi sebagai usaha *Brand Awareness* agar pelanggan yang pasif bisa kembali melakukan transaksi pembelian paket *coaching* pada perusahaan AEGIS (Puspitasari et al., 2020).

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

BAB 5
KESIMPULAN DAN SARAN

5.1 Kesimpulan

Tahap akhir dari penelitian ini dapat disimpulkan hasil penelitian "Analisis Loyalitas Pelanggan AEGIS Pada Pembelian Paket Coaching Menggunakan Algoritma Bisecting K-Means" bahwa:

1. Hasil penelitian algoritma *Bisecting K-Means* dapat diimplementasikan pada data pelanggan AEGIS dengan menggunakan 3 *cluster* berdasarkan kategori loyalitas pelanggan yaitu, pelangan loyal, moderat dan tidak loyal. Selain itu, Hasil *Silhouette Coefficient* didapatkan sebesar 0,81 pada 2 *cluster* dan 0,45 pada 3 cluster. Berdasarkan hasil *score* pada setiap *cluster*, 2 *cluster* memiliki *score* paling besar dan dapat dikatakan sebagai *cluster* kuat karena data sudah berada pada *cluster* yang benar. Sedangkan, 3 *cluster* dikatakan sebagai *cluster* lemah karena masih ada data yang tidak berada pada *cluster* sebenarnya.
2. Berdasarkan hasil perhitungan persentase loyalitas pelanggan, *cluster* 2 merupakan kelompok yang loyal karena memiliki jumlah persentase yang tinggi pada pembelian paket *coaching.* Pelanggan paling banyak memiliki MMR antara 2001 sampai 4000 dan paling banyak membeli paket dengan 5 sesi. Kemudian, *cluster* 1 dikatakan sebagai kelompok yang moderat karena memiliki jumlah persentase loyalitas terbesar kedua. Pada *cluster* 1, rata-rata pelanggan yang membeli paket *coaching* memiliki MMR antara 2001 sampai 6000. Selain itu, jumlah pembelian sesi paling banyak pada 1 sesi. *Cluster* 3 dikatakan sebagai kelompok yang tidak loyal karena memiliki jumlah pendaftaran paling sedikit. Strategi CRM yang bisa dilakukan oleh perusahaan AEGIS untuk mempertahankan pelanggan lama maupun memperoleh pelanggan baru dengan memberikan diskon dan kartu member.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

5.2 Saran

Berdasarkan hasil penelitian"Analisis Loyalitas Pelanggan AEGIS Pada Pembelian Paket *Coaching* Menggunakan Algoritma *Bisecting K-Means*" terdapat kekurangan yang dilakukan sehingga peneliti berharap pada penelitian kedepan peneliti memberikan beberapa saran sebagai berikut:

1. Penelitian ini bisa dijadikan acuan untuk penelitian selanjutnya menggunakan metode yang sama ataupun berbeda dengan data yang lebih banyak pada masa mendatang. Harapan peneliti hasil penelitian selanjutnya memiliki hasil akurasi yang lebih tinggi.
2. Data pelanggan AEGIS yang digunakan lebih luas yang menyangkut game game lain untuk dilakukan pengolahan data pada penelitian kedepan.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

DAFTAR PUSTAKA

Abirami, K., & Mayilvahanan, P. (2016). Performance Analysis of K-Means and Bisecting K-Means Algorithms in Weblog Data. *International Journal of Emerging Technologies in Engineering Research (IJETER)*, *4*(8), 119–124. www.ijeter.everscience.org

Anggara, M., Sujiani, H., & Helfi, N. (2016). Pemilihan Distance Measure Pada K-Means Clustering Untuk Pengelompokkan Member Di Alvaro Fitness. *Jurnal Sistem Dan Teknologi Informasi*, *1*(1), 1–6.

Anitha, P., & Patil, M. M. (2020). RFM model for customer purchase behavior using K-Means algorithm. *Journal of King Saud University - Computer and Information Sciences*. https://doi.org/10.1016/j.jksuci.2019.12.011

Asroni, R. A. (2015). Penerapan Metode K-Means Untuk Clustering Mahasiswa Berdasarkan Nilai Akademik Dengan Weka Interface Studi Kasus Pada Jurusan Teknik Informatika UMM Magelang. *Ilmiah Semesta Teknika*, *18*(1), 76–82. https://doi.org/10.1038/hdy.2009.180

Azuri, D. F., Zulhanif, & Pontoh, R. S. (2016). Pengelompokkan Kabupaten/Kota Di Pulau Jawa Berdasarkan Pembangunan Manusia Berbasis Gender Menggunakan Bisecting K-Means. *Peran Penelitian Ilmu Dasar Dalam Menunjang Pembangunan Berkelanjutan*, 78–83. http://riset.fmipa.unpad.ac.id/proc_semnasmipa2016/017. 078 - 083 Dila Fitriani.pdf

Banu, S., & C.R, A. (2017). Tweet Clustering Using Bisecting K-Means. *International Journal of Innovative Research in Computer and Communication Engineering*, *5*(3), 5194–5202. https://doi.org/10.15680/IJIRCCE.2017

Dewi, N. V. U. (2019). *Hubungan Kerja Antara Atlit E-sports Dengan Pengusaha CV Pember Kerja*. *2*(1), 256–274.

Funk, D. C., Pizzo, A. D., & Baker, B. J. (2018). eSport management: Embracing eSport education and research opportunities. *Sport Management Review*,

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

21(1), 7–13. https://doi.org/10.1016/j.smr.2017.07.008

Gustientiedina, G., Adiya, M. H., & Desnelita, Y. (2019). Penerapan Algoritma K-Means Untuk Clustering Data Obat-Obatan. *Jurnal Nasional Teknologi Dan Sistem Informasi*, *5*(1), 17–24. https://doi.org/10.25077/teknosi.v5i1.2019.17-24

Halim, A., Gohzali, H., Panjaitan, D. M., & Maulana, I. (2017). Sistem Rekomendasi Film menggunakan Bisecting K-Means dan Collaborative Filtering. *Citisee*, *1*(3), 37–41.

Hardiani, T., Sulistyo, S., & Hartanto, R. (2014). Kajian Data Mining Customer Relationship Management pada Lembaga Keuangan Mikro. *Jurnal Sistem Informasi Bisnis*, *4*(3), 181–187. https://doi.org/10.21456/vol4iss3pp181-186

Hemeida, A. M., Alkhalaf, S., Mady, A., Mahmoud, E. A., Hussein, M. E., & Baha Eldin, A. M. (2019). Implementation of nature-inspired optimization algorithms in some data mining tasks. *Ain Shams Engineering Journal*, *xxxx*. https://doi.org/10.1016/j.asej.2019.10.003

Hosseini, S. M. S., Maleki, A., & Gholamian, M. R. (2010). Cluster analysis using data mining approach to develop CRM methodology to assess the customer loyalty. *Expert Systems with Applications*, *37*(7), 5259–5264. https://doi.org/10.1016/j.eswa.2009.12.070

Jaskolka, K., Seiler, J., Beyer, F., & Kaup, A. (2019). A Python-based laboratory course for image and video signal processing on embedded systems. *Heliyon*, *5*(10). https://doi.org/10.1016/j.heliyon.2019.e02560

Joyendri, A. (2017). Strategi Customer Relationship Management Untuk Meningkatkan Loyalitas Pelanggan Dan Volume Penjualan Menggunakan Teknik Clustering K-Means. *Telematika*, *14*(2), 75. https://doi.org/10.31315/telematika.v14i2.2094

Kamat, P. V., Johnson, P. T., Chodankar, R., Harmalkar, R., Naik, G., & Narulkar, P. (2017). Document Clustering Using Divisive Hierarchical Bisecting Min Max Clustering Algorithm. *IOSR Journal of Computer Engineering*, *19*(03), 66–70. https://doi.org/10.9790/0661-1903066670

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Lubis, A. H. (2016). Model Segmentasi Pelanggan Dengan Kernel K-Means Clustering Berbasis Customer Relationship Management. *Jurnal & Penelitian Teknik Informatika*, *1*, 36–41.

Lutz, M. (2010). *Programming Python* (July Steele (ed.); Fourth). O'Reilly Media, Inc.

Macedo, T., & Falcão, T. (2020). Like a Pro: Communication, Camaraderie and Group Cohesion in the Amazonian Esports Scenario. *Entertainment Computing*, 100354. https://doi.org/10.1016/j.entcom.2020.100354

Macey, J., & Hamari, J. (2018). Investigating Relationships Between Video Gaming, Spectating Esports, and Gambling. *Computers in Human Behavior*, *80*, 344–353. https://doi.org/10.1016/j.chb.2017.11.027

Majhi, S. K., & Biswal, S. (2018). Optimal cluster analysis using hybrid K-Means and Ant Lion Optimizer. *Karbala International Journal of Modern Science*, *4*(4), 347–360. https://doi.org/10.1016/j.kijoms.2018.09.001

Mardi, Y. (2019). Data Mining : Klasifikasi Menggunakan Algoritma C4.5. *Jurnal Edik Informatika*, *2*(2), 213–219.

Matthes, E. (2016). *Python Cash Course* (Hoffman Riley (ed.)). William Pollock.

Maulida, L. (2018). Penerapan Datamining Dalam Mengelompokkan Kunjungan Wisatawan Ke Objek Wisata Unggulan Di Prov. Dki Jakarta Dengan K-Means. *JISKA (Jurnal Informatika Sunan Kalijaga)*, *2*(3), 167. https://doi.org/10.14421/jiska.2018.23-06

Meng, T., Jing, X., Yan, Z., & Pedrycz, W. (2020). A survey on machine learning for data fusion. *Information Fusion*, *57*(2), 115–129. https://doi.org/10.1016/j.inffus.2019.12.001

Nishom, M. (2019). Perbandingan Akurasi Euclidean Distance, Minkowski Distance, dan Manhattan Distance pada Algoritma K-Means Clustering berbasis Chi-Square. *Jurnal Informatika: Jurnal Pengembangan IT*, *4*(1), 20–24. https://doi.org/10.30591/jpit.v4i1.1253

Patil, R., & Khan, A. (2015). Bisecting K-Means for Clustering Web Log data.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

International Journal of Computer Applications, *116*(19), 36–41. https://doi.org/10.5120/20448-2799

Pedregosa, F., Weiss, R., & Brucher, M. (2011). Scikit-learn : Machine Learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Pierre Carbonnelle. (2019). *PYPL PopularitY of Programming Language index*. 2020. http://pypl.github.io/PYPL.html

Pratama, A., Rahmad, S., & Mahfuz, S. (2019). *Aplikasi Data Mining Dengan Menggunakan Metode K-Means Untuk Mengetahui Pola Loyalitas Pelanggan Pada*.

Puspitasari, N., Widians, J. A., & Setiawan, N. B. (2020). Segmentasi pelanggan menggunakan algoritme bisecting k-means berdasarkan model recency, frequency, dan monetary (RFM). *Jurnal Teknologi Dan Sistem Komputer*, *8*(2), 78–83. https://doi.org/10.14710/jtsiskom.8.2.2020.78-83

Qiao, J., Cai, X., Xiao, Q., Chen, Z., Kulkarni, P., Ferris, C., Kamarthi, S., & Sridhar, S. (2019). Data on MRI brain lesion segmentation using K-means and Gaussian Mixture Model-Expectation Maximization. *Data in Brief*, *27*, 104628. https://doi.org/10.1016/j.dib.2019.104628

Rachman, J. B., Adityani, S., Suryadipura, D., & Utama, B. P. (2020). Esport Sebagai Sumber Soft Power Indonesia : Sosialisasi Kepada Anak Muda. *CARADDE: Jurnal Pengabdian Kepada Masyarakat*, *3*(1), 43–52. https://journal.ilininstitute.com/index.php/caradde

Rahman, A. T., Wiranto, & Rini, A. (2017). Coal Trade Data Clustering Using K-Means (Case Study Pt. Global Bangkit Utama). *ITSMART: Jurnal Teknologi Dan Informasi*, *6*(1), 24–31. https://jurnal.uns.ac.id/itsmart/article/download/11296/11108

Ramdan, E., Samad, A., & Khaeruddin. (2016). Jurnal Pendidikan Fisika Universitas Muhammadiyah Makassar Meningkatkan Hasil Belajar Fisika Peserta Didik Melalui Strategi. *Jurnal Pendidikan Fisika Universitas Muhammadiyah Makassar*, *5*, 15–24.

Safitri, A. Y., Wibawa, B. M., & Ardiantono, D. S. (2019). Pengukuran Kepuasan

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

dan Loyalitas Konsumen terhadap Kualitas Layanan Taksi Resmi Bandar Udara Juanda Surabaya. *Jurnal Sains Dan Seni ITS*, *8*(1). https://doi.org/10.12962/j23373520.v8i1.41573

Satish, L., & Yusof, N. (2017). A Review: Big Data Analytics for enhanced Customer Experiences with Crowd Sourcing. *Procedia Computer Science*, *116*, 274–283. https://doi.org/10.1016/j.procs.2017.10.058

Saylı, A., Ozturk, I., & Ustunel, M. (2016). Brand loyalty analysis system using K-Means algorithm. *Journal of Engineering Technology and Applied Sciences*, *1*(3), 107–126. https://doi.org/10.30931/jetas.287786

Shang, C., & You, F. (2019). Data Analytics and Machine Learning for Smart Process Manufacturing: Recent Advances and Perspectives in the Big Data Era. *Engineering*, *5*(6), 1010–1016. https://doi.org/10.1016/j.eng.2019.01.019

Sharma, A., Jit, H., & Sidhu, S. (2018). *Customer Relationship Management Using Clustering And Classification Technique*. *20*(5), 67–72. https://doi.org/10.9790/0661-2005026772

Sholehhudin, M., Fauzi Ali, M., & Adinugroho, S. (2018). *Implementasi Metode Text Mining dan K-Means Clustering untuk Pengelompokan Dokumen Skripsi (Studi Kasus : Universitas Brawijaya)*. *2*(11), 5518–5524.

Suryani, M. A. I., Arifin, Z., & Hatta, H. R. (2017). Pemilihan Paket Wisata Menggunakan Metode Analytical Hierarchy Process (AHP). *Informatika Mulawarman : Jurnal Ilmiah Ilmu Komputer*, *12*(2), 64. https://doi.org/10.30872/jim.v12i2.646

Swapna, K., & Babu, P. M. S. P. (2017). A Framework for Outlier Detection Using Improved Bisecting k-Means Clustering Algorithm. *International Journal of Electrical & Computer Sciences IJECS-IJENS*, *17*(April), 2.

Syakur, M. A., Khotimah, B. K., Rochman, E. M. S., & Satoto, B. D. (2018). Integration K-Means Clustering Method and Elbow Method for Identification of the Best Customer Profile Cluster. *IOP Conference Series: Materials Science and Engineering*, *336*(1). https://doi.org/10.1088/1757-899X/336/1/012017

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Umar, R., Fadlil, A., & Yuminah, Y. (2018). Sistem Pendukung Keputusan dengan Metode AHP untuk Penilaian Kompetensi Soft Skill Karyawan. *Khazanah Informatika: Jurnal Ilmu Komputer Dan Informatika*, *4*(1), 27. https://doi.org/10.23917/khif.v4i1.5978

Wahyuni, I., Auliya, Y. A., & Mahmudy, W. F. (2016). Clustering Nasabah Bank Berdasarkan Tingkat Likuiditas Menggunakan Hybrid PSO K-Means. *Jitika*, *10*(1).

Waring, J., Lindvall, C., & Umeton, R. (2020). Automated Machine Learning: Review of the State-of-the-Art and Opportunities for Healthcare. *Artificial Intelligence in Medicine*, 101822. https://doi.org/10.1016/j.artmed.2020.101822

Wiguna, M. abdilah, & Wawan, ardiyan s. (2016). Bandung eSport Arena Membangun Ekosistem eSport di Indonesia. *Jurnal Sains Dan Seni Its*, *5*(2), 39–42.

Zhao, Y., Zhang, C., Zhang, Y., Wang, Z., & Li, J. (2019). A review of data mining technologies in building energy systems: Load prediction, pattern identification, fault detection and diagnosis. *Energy and Built Environment*. https://doi.org/10.1016/j.enbenv.2019.11.003

Zhuang, Y., Mao, Y., & Chen, X. (2016). A Limited-Iteration Bisecting K-means for Fast Clustering Large Datasets. *IEEE Computer Society*, 2259–2264. https://doi.org/10.1109/TrustCom.2016.346

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id