

BAB II

TINJAUAN PUSTAKA DAN DASAR TEORI

2.1. Tinjauan Pustaka

Klasifikasi produk atau paket merupakan tugas yang biasa dilakukan dalam riset pasar, khususnya untuk audit pelacakan penjualan dan layanan terkait. Sumber data elektronik telah menyebabkan peningkatan volume, baik dalam volume penjualan yang dilacak dan juga jumlah paket (unit penyimpanan stok). Bertambahnya paket menimbulkan masalah pada saat klasifikasi produk. Klasifikasi menggunakan cara manual menimbulkan biaya yang besar, karena memerlukan orang banyak untuk memproses klasifikasi secara akurat dan tepat waktu. Mengurangi metode konvensional adalah prioritas utama untuk audit berbasis bisnis riset pasar, yang mengarah pada otomatisasi melalui metode *machine learning*. Dengan menerapkan metode klasifikasi hirarkis untuk mengambil manfaat dari struktur produk yang diklasifikasi. Setelah model dilatih, mereka dapat digunakan pada data yang tidak diklasifikasikan. Harapannya metode dapat belajar untuk mengklasifikasikan secara akurat sehingga pekerjaan manual dikurangi hingga level yang dapat dikelola (Hang dan Banks, 2019).

Penelitian dilakukan dengan menggunakan *machine learning* dalam sistem berbasis aturan untuk menyaring pesan yang tampil dalam media *online*. Sistem berbasis aturan menggunakan *machine learning* memberikan kemampuan kepada pengguna *On-line Social Networks* (OSNs) dalam mengatur pesan yang diposting di ruang pribadi supaya konten yang ditampilkan tidak berlebihan. Aturan yang digunakan dalam penyaringan pesan dalam penelitian berdasarkan tiga masalah utama yaitu masalah kehidupan sehari-hari, kemudian pesan yang memiliki arti yang sama, dan hubungan berdasarkan pada siapa yang menulis pesan. Aturan yang dibuat dalam penyaringan memungkinkan pengguna untuk membuat batasan pada pesan berdasarkan beberapa kriteria yang berbeda, kemudian *machine learning* mengklasifikasikan secara otomatis dengan memberikan label pada pesan untuk mendukung penyaringan berdasarkan konten (Mehdia dkk., 2014).

Penelitian lain membahas pemanfaatan *machine learning* berbasis *rule induction* untuk mendeteksi autism. Gangguan spektrum *autisme* adalah gangguan

perkembangan yang menggambarkan tantangan tertentu yang berkaitan dengan komunikasi keterampilan sosial (verbal dan non-verbal) dan perilaku repetitif. Untuk meningkatkan akurasi dan efisiensi dalam mendeteksi gangguan spektrum *autisme* dengan membangun sistem klasifikasi menggunakan teknologi cerdas seperti *Machine Learning*. Di dalam artikel tersebut diusulkan metode *machine learning* baru yang disebut *Rules-Machine Learning* (RML) yang tidak hanya mendeteksi sifat autisme dari kasus dan kontrol tetapi juga menawarkan basis pengetahuan pengguna (aturan) yang dapat digunakan oleh para ahli domain dalam memahami alasan di balik klasifikasi. Hasil empiris pada tiga set data yang terkait dengan anak-anak, remaja, dan orang dewasa menunjukkan bahwa *Rules-Machine Learning* menawarkan pengklasifikasi dengan akurasi prediksi, sensitivitas, rata-rata harmonis, dan spesifisitas yang lebih tinggi daripada pendekatan *machine learning* lainnya seperti *Boosting*, *Bagging*, pohon keputusan, dan aturan induksi (Thabtah dan Peebles, 2019).

Penelitian konstruksi jaringan fungsional di *Arabidopsis* menggunakan *rule-based machine learning* pada kumpulan data skala besar. Dalam penelitiannya, disajikan metodologi komputasi menggunakan *rule-based machine learning* untuk mengidentifikasi hubungan fungsional antara gen dari data set *microarray*. Penelitian menggunakan pendekatan yang disebut "*co-prediction*," berdasarkan pada aturan kemampuan kolektif kelompok gen yang terjadi bersamaan untuk memprediksi hasil perkembangan sistem biologis secara akurat. Penelitian menunjukkan fungsi dari *co-prediction* sebagai alat analitik yang efektif dengan menggunakan data *microarray* yang tersedia untuk umum yang dihasilkan secara eksklusif dari biji *Arabidopsis thaliana* untuk menghitung jaringan interaksi gen fungsional, disebut *Seed Co-Prediction Network* (SCoPNet). SCoPNet memprediksi hubungan fungsional antara gen yang bertindak dalam jalur transduksi sinyal dan perkembangan yang sama terlepas dari kesamaan dalam pola ekspresi masing-masing gen (Bassel dkk., 2011).

Proyek MOSAIC yang dibiayai oleh Uni Eropa telah menambahkan algoritma *machine learning* ke dalam aliran data mining, dengan menggabungkannya dengan statistik klasik untuk mengekstraksi pengetahuan dari

data. Aliran *data mining* digunakan untuk mendapatkan serangkaian model prediksi *type 2 diabetes mellitus* (T2DM) berdasarkan data catatan. Aliran *data mining* terdiri dari profil pusat klinis, penargetan model prediktif, konstruksi model prediktif, dan validasi model. Setelah berhubungan dengan data yang hilang melalui *random forest* (RF) dan menggunakan strategi yang sesuai untuk menangani ketidakseimbangan kelas, digunakan *regresi logistik* dengan pemilihan fitur bertahap untuk memprediksi timbulnya *retinopati*, *neuropati*, atau *nefropati*, sejak kunjungan pertama di Rumah Sakit Pusat Diabetes (bukan dari diagnosis) pada skenario waktu yang berbeda, pada 3, 5, dan 7 tahun. Faktor yang dipertimbangkan adalah usia, jenis kelamin, waktu dari diagnosis, *body mass index* (BMI), *glycated hemoglobin* (HbA1c), tekanan darah tinggi, dan kebiasaan merokok. Variabel yang berbeda dipilih untuk setiap skenario komplikasi dan waktu, yang mengarah ke model khusus yang mudah diterjemahkan ke praktik klinis. Hasil dari penelitian menunjukkan *data mining* dan metode komputasi dapat secara efektif diadopsi dalam ilmu kedokteran untuk mendapatkan model yang menggunakan informasi spesifik pasien untuk memprediksi hasil yang diinginkan (Dagliati dkk., 2018).

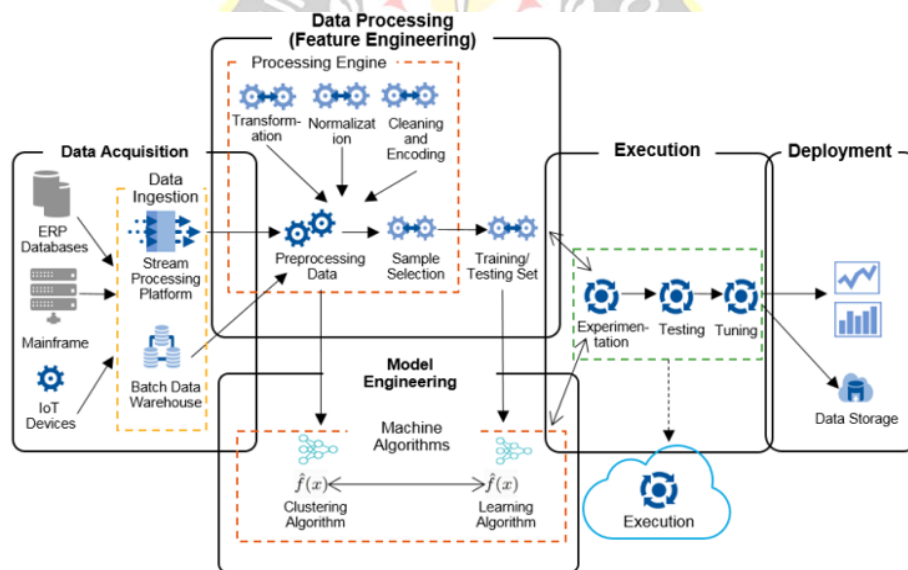
Metodologi *machine learning* digunakan dalam memprediksi keberhasilan atau kegagalan peserta didik dalam MOOCs (Mourdi dkk., 2019). MOOCs (Massive Open Online Courses) merupakan sistem pembelajaran jarak jauh yang sudah menjadi tren. Dalam pelatihan MOOCs banyak peserta yang tidak menyelesaikan pelatihannya, dari banyaknya peserta yang terdaftar rata-rata hanya 10 persen saja yang berhasil selesai dan mendapatkan sertifikat. Penelitian ini mengajukan metodologi untuk memprediksi tentang perilaku peserta didik dengan menggunakan data *Stanford* yang dibagi dalam beberapa tahap yaitu : ekstraksi data, studi eksplorasi dan kemudian analisis multivariat untuk mengurangi dimensi dan mengekstraksi fitur yang paling relevan. Langkah kedua adalah perbandingan antara lima algoritma *machine learning*. Langkah yang terakhir menggunakan prinsip aturan asosiasi untuk mengekstrak kesamaan antara perilaku peserta didik yang keluar dari MOOCs. Hasil dari penelitian menunjukkan deep learning memberikan tingkat prediksi yang relatif tinggi (Mehdia dkk., 2014).

2.2. Dasar Teori

2.2.1 Machine Learning

Machine learning adalah salah satu bidang di mana kecerdasan buatan yang mengembangkan algoritma yang dapat mempelajari pola dan aturan keputusan dari data (Seffens dkk., 2015). *Machine learning* berfokus pada pengembangan program komputer yang dapat mengakses data dan menggunakannya untuk belajar sendiri.

Infrastruktur aplikasi *machine learning* sangat fleksibel disesuaikan dengan kebutuhan skala proses dan volume data yang proses. Arsitektur *machine learning* diilustrasikan pada Gambar 2.1.



Gambar 2.1. Arsitektur *machine learning* (Sapp, 2017).

Gambar 2.1 menunjukkan ilustrasi arsitektur dari *machine learning* yang terdiri dari *Data acquisition*, *Data processing*, *Data modeling or model engineering*, *Execution*, *Deployment*. Penjelasan dari masing-masing bagian adalah sebagai berikut :

- Pengambilan data (*Data acquisition*).

Data dikumpulkan, disiapkan dan kemudian diteruskan untuk diproses

- Pemrosesan data (*Data processing*).

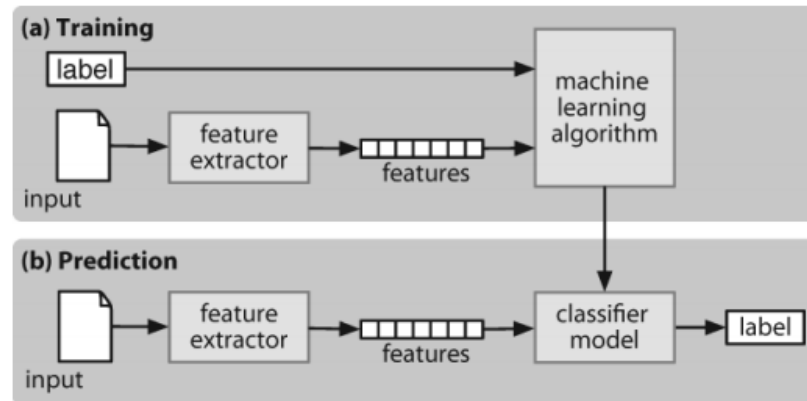
Langkah-langkah seperti preprocessing, pemilihan sampel dan pelatihan dataset berlangsung, dalam persiapan untuk pelaksanaan *machine learning*. Analisis

fitur atau rekayasa fitur (*Feature analysis or feature engineering*) (bagian dari komponen pemrosesan data), fitur yang menggambarkan struktur yang melekat dalam data dipilih dan dianalisis.

- c. *Pemodelan data atau rekayasa model (Data modeling or model engineering)*.
Desain model data dan algoritma yang digunakan dalam pemrosesan data ML (termasuk algoritma pengelompokan dan pelatihan):
 1. *Model fitting* (di mana satu set data pelatihan ditugaskan ke model untuk membuat prediksi yang dapat diandalkan pada data baru atau tidak terlatih)
 2. *Evaluasi model (Model evaluation)*. Model dievaluasi berdasarkan kinerja dan kemanjuran
- d. *Eksekusi (Execution)*.
Lingkungan di mana data yang diproses, dan dilatih diteruskan untuk digunakan dalam pelaksanaan rutinitas ML (seperti eksperimen, pengujian, dan penyetalan)
- e. *Penempatan (Deployment)*.
Pemanfaatan dari ML - seperti model atau wawasan - digunakan untuk aplikasi perusahaan, sistem atau penyimpanan data (misalnya, untuk pelaporan)

Proses pembelajaran dimulai dengan pengamatan atau data, seperti contoh, pengalaman langsung, atau instruksi, untuk mencari pola dalam data dan membuat keputusan yang lebih baik dimasa depan berdasarkan contoh yang diberikan. Tujuan utamanya adalah untuk memungkinkan komputer belajar secara otomatis tanpa campur tangan atau bantuan manusia dan menyesuaikan tindakan yang sesuai. Beberapa metode *Machine Learning* :

- a) *Supervised Machine Learning* (Chen dkk., 2015), algoritma mengambil satu set contoh sebagai input (disebut contoh pelatihan), kemudian setiap pelatihan kelas memiliki label tertentu yang bisa dikenali dan memiliki serangkaian fitur terkait dengan nilai-nilai. Model kemudian menampilkan atau memprediksi kelas contoh baru dengan memberikan nilai fitur, hasilnya dikaitkan dengan sensitivitas tertentu (disebut akurasi dalam ilmu informasi) (Awaysheh dkk., 2016). Algoritma *supervised machine learning* dengan metode klasifikasi di tunjukkan pada Gambar 2.2.



Gambar 2.2. *Supervised Classification* (Bird dkk., 2009)

Gambar 2.2 menjelaskan selama pelatihan, fitur ekstraktor digunakan untuk mengkonversi setiap nilai input ke set fitur. Set fitur ini, yang menangkap informasi dasar tentang setiap input yang harus digunakan untuk mengklasifikasikannya, pasangan set fitur dan label dimasukkan ke dalam algoritma pembelajaran mesin untuk menghasilkan model. Selama prediksi, ekstraktor fitur yang sama digunakan untuk mengkonversi input yang tidak diketahui ke set fitur. Set fitur ini kemudian dimasukkan ke dalam model, kemudian menghasilkan label yang diprediksi (Bird dkk., 2009).

- b) *Unsupervised Machine Learning*, mempelajari bagaimana sistem pembelajaran tanpa pengawasan dapat menyimpulkan suatu fungsi untuk menggambarkan struktur tersembunyi dari data yang tidak berlabel. Alasan untuk menggunakan alat jenis ini mungkin karena kelas untuk *instance* tidak diketahui, data *historis* untuk pelatihan algoritma tidak tersedia, atau pengguna mungkin ingin menjelajahi klasifikasi baru untuk data tersebut (Awaysheh dkk., 2019). Sistem tidak mengetahui *output* yang tepat, tetapi mengeksplorasi data dan dapat menarik kesimpulan dari dataset untuk menggambarkan struktur tersembunyi dari data yang tidak berlabel.
- c) *Reinforcement Machine Learning* (Awaysheh dkk., 2016), *Machine Learning* di mana agen mempelajari sesuatu hal dengan cara melakukan aksi tertentu dan melihat hasil dari aksi tersebut (belajar berdasarkan pengalaman yang dialami oleh agen tersebut). Jadi *Reinforcement Learning* didasarkan pada interaksi agen yang mengeksekusi tindakan dan lingkungannya yang memberikan umpan

balik positif atau negatif, misalnya hadiah. Pada *Reinforcement Learning* (RL), proses belajar dapat digambarkan sebagai sebuah loop di mana:

- 1) Agen menerima sebuah *State* (S0) dari *Environment* yang ada.
- 2) Berdasarkan *State* tersebut (S0), agen akan melakukan sebuah aksi (A0).
- 3) *Environment* akan berubah menjadi *State* baru (S1).
- 4) *Environment* akan memberikan sebuah '*Reward*' (R1) pada agen.

Proses tersebut akan dilakukan berulang, yang tujuannya agen tersebut dapat memaksimalkan kumulatif reward yang didapat.

Machine learning memungkinkan analisis data dalam jumlah besar. *Machine learning* umumnya memberikan hasil yang lebih cepat, lebih akurat untuk mengidentifikasi peluang yang menguntungkan atau risiko berbahaya, mungkin juga memerlukan waktu dan sumber daya tambahan untuk melatih *machine learning* dengan benar. *Machine learning* dengan digabungkan dengan teknologi kognitif dapat membuat pekerjaan lebih efektif dalam memproses informasi dalam volume besar.

2.2.2 Sistem berbasis aturan (*Rule-based system*).

Sistem berbasis aturan memiliki basis pengetahuan yang direpresentasikan sebagai kumpulan "aturan" yang biasanya dinyatakan sebagai klausa "jika-maka" (Kovarik, 2009). Seperangkat aturan membentuk basis pengetahuan yang diterapkan pada serangkaian fakta saat ini. Sistem berbasis aturan terdiri dari "if-then" aturan, sekelompok data dan penerjemah sebagai pengendali dari aturan yang diterapkan dalam memberikan fakta-fakta (Anugrah dkk, 2018).

Metode yang digunakan oleh sistem berbasis aturan dalam merepresentasikan pengetahuan inferensial dengan menggunakan bentuk sederhana "if-then", yang relatif mudah untuk dinyatakan dan dipahami sehingga paradigma aturan secara alami mudah untuk dipahami oleh manusia. Pembelajaran dengan aturan memiliki beberapa keuntungan sebagai berikut.

1. Orang dapat dengan mudah memahami dan menafsirkannya karena aturan sangat alami untuk representasi pengetahuan.

2. Berdasarkan pada basis data aturan dan input data dari pengguna, sehingga dapat dijelaskan aturan atau kumpulan aturan mana yang digunakan dalam menyimpulkan label kelas sehingga pengguna jelas tentang logika di balik inferensi.
3. Model klasifikasi berbasis aturan dapat dengan mudah ditingkatkan dan dilengkapi dengan menambahkan aturan baru dari para ahli berdasarkan pengetahuan bidangnya.
4. Setelah aturan dipelajari dan disimpan ke dalam basis data aturan, selanjutnya dapat menggunakannya untuk mengklasifikasikan contoh baru dengan cepat melalui membangun struktur indeks untuk aturan dan mencari aturan yang relevan secara efisien.

Aturan dapat mewakili informasi atau pengetahuan dengan cara yang sangat sederhana dan efektif. Sistem berbasis aturan menyediakan model data yang sangat baik yang bisa dipahami manusia dengan sangat baik. Aturan disajikan dalam bentuk logika seperti pernyataan *if-then*. Aturan yang umum digunakan dapat dinyatakan sebagai berikut:

If kondisi *Then* kesimpulan

Bagian *if* disebut "anteseden" atau "kondisi" dan bagian *Then* disebut "konsekuen" atau "kesimpulan". Pada dasarnya jika kondisi aturan terpenuhi, dapat menyimpulkan atau mengurangi kesimpulan.

Suatu aturan dapat memiliki beberapa kondisi yang bisa dihubungkan dengan dengan kata kunci *AND*, *OR*, atau kombinasi dari keduanya. kondisi suatu aturan terdiri dari dua bagian, objek dan nilainya. Objek dan nilai dihubungkan oleh operator. Operator dapat bersifat matematis atau mungkin logika.

Contoh dari basis aturan yang terdiri dari set aturan

Aturan 1: *If A and C then F*

Aturan 2: *If A or E then G*

Aturan 3: *If B then E*

2.2.3 Sistem Pakar

Sistem pakar merupakan bagian dari kecerdasan buatan yang terdiri dari pengetahuan dan pengalaman dari banyak pakar yang dimasukkan ke dalam suatu basis pengetahuan. Sistem pakar dapat membantu seseorang yang mungkin bukanlah seorang pakar untuk menyelesaikan persoalan tertentu sesuai dengan basis pengetahuan yang telah dimasukkan ke dalam sistem. Dengan adanya basis pengetahuan yang setiap saat dapat bertambah, maka sistem pakar akan semakin berkembang keakuratannya (Solichin, 2011).

Sistem pakar terdiri dari beberapa konsep yang harus dimilikinya (Solichin, 2011). Konsep dasar dari suatu sistem pakar antar lain :

1. Keahlian.

Suatu pengetahuan khusus yang diperoleh dari latihan, belajar, dan pengalaman. Pengetahuan dapat berupa fakta, teori, aturan, strategi global untuk memecahkan masalah.

2. Ahli (*Expert*).

Ahli harus mampu melakukan kegiatan mengenali dan memformulasikan permasalahan, memecahkan masalah secara cepat dan tepat, menerangkan pemecahannya, belajar dari pengalaman, merestrukturisasi pengetahuan, memecahkan aturan serta menentukan relevansi.

3. Memindahkan kemampuan (*Transferring Expertise*).

Proses pentransferan keahlian dari seorang pakar ke dalam komputer agar dapat digunakan oleh orang lain yang bukan pakar. Pengetahuan tersebut ditempatkan ke dalam sebuah komponen yang dinamakan basis pengetahuan

4. Menyimpulkan aturan (*Inferencing Rule*).

Inferencing Rule merupakan kemampuan komputer yang telah diprogram. Penyimpulan ini dilakukan oleh mesin inferensi yang meliputi prosedur tentang menyelesaikan masalah

5. Peraturan (*Rule*).

Peraturan diperlukan karena mayoritas dari sistem pakar bersifat *rule-based systems*, yang berarti pengetahuan disimpan dalam bentuk peraturan.

6. Kemampuan dalam menjelaskan (*Explanation Capability*).

Karakteristik dari sistem pakar yang memiliki kemampuan menjelaskan atau memberi saran mengapa tindakan tertentu dianjurkan atau tidak dianjurkan.

2.2.4 Perilaku Bermasalah

Perilaku manusia merupakan ungkapan atau ekspresi dari ciri yang mewarnai atau melandasi individu itu berperilaku sehingga perilaku tersebut dapat diobservasi. Observasi atau pengamatan untuk mengumpulkan data tentang perilaku (For, 2012), bermanfaat untuk memahami diri dari seseorang sehingga bisa berguna memahami permasalahan yang sedang dialami oleh seseorang.

Peserta didik ketika menginjak usia remaja di mana mulai mengenal lingkungan atau masyarakat yang lebih luas yang selalu dihadapkan pada permasalahan-permasalahan yang lebih rumit yang memerlukan penanganan yang sangat serius. Tanggapan guru tentang keterampilan siswa dan perilaku bermasalah sangat berhubungan dengan hasil dari akademik (Wanzek dkk., 2014).

Perilaku bermasalah dapat diartikan segala tingkah laku yang mengganggu dalam proses belajar siswa (Johnson dan Hannon, 2014). Perilaku bermasalah dalam proses pembelajaran di sekolah dapat menjadikan motivasi, perhatian, dan semangat menjadi lebih rendah (Bulotsky-Shearer dkk., 2011). Dalam studi yang sama, perilaku bermasalah siswa dapat mendasarinya siswa kurang berprestasi dan hilangnya kontrol diri (Bulotsky-Shearer dkk., 2011).

Masalah peserta didik merupakan hal yang harus diketahui oleh konselor, penyebab masalah ini yang mungkin menjadikan masalah di kemudian hari. Dalam mengidentifikasi masalah peserta didik banyak cara yang bisa dilakukan salah satunya adalah dengan Daftar Cek Masalah. Dalam daftar cek masalah berisi sejumlah kemungkinan masalah yang pernah atau sedang dihadapi oleh individu.

Daftar cek masalah berisi 240 butir pernyataan masalah yang terbagi dalam 12 bidang kategori antara lain :

1. Kesehatan.
2. Keadaan Ekonomi.
3. Kehidupan Keluarga.
4. Agama Dan Moral.
5. Kegemaran.
6. Hubungan Pribadi.
7. Kehidupan Sosial Keaktifan Berorganisasi.
8. Masalah Remaja.
9. Penyesuaian Terhadap Sekolah.
10. Penyesuaian Terhadap Kurikulum.
11. Kebiasaan Belajar.
12. Masa Depan Dan Cita - Cita Pendidikan/ Jabatan.

a) Langkah dalam menganalisis dalam identifikasi masalah peserta didik :

1. Menjumlahkan butir item yang menjadi masalah pada individu.

$$\sum_{i=1}^m Ni = N_1+N_2+N_3\dots+N_m$$

2. Menggunakan persentase per-topik masalah, dengan cara mencari rasio antara butir masalah yang menjadi masalah dengan jumlah butir topik masalah dengan menggunakan rumus :

$$X = \frac{\sum_{i=1}^m Ni}{Nm} \times 100\%$$

Keterangan :

X : Persentase per-topik masalah

N : Jumlah item yang menjadi masalah pada satu topik masalah.

Nm : Jumlah item pada topik masalah.

3. Mencari jenjang masalah dengan cara mengurutkan persentase dari topik masalah.
4. Mengkonversikan persentase masalah ke standar skala, seperti yang ditunjukkan pada Tabel 2.1.

Tabel 2.1. Skala permasalahan.

Skala Persentase %	Predikat	Keterangan
0	A (Baik Sekali)	Tanpa Masalah
1 - 10	B (Baik)	Sedikit bermasalah
11 - 25	C (Sedang)	Cukup bermasalah
26 – 50	D (Kurang)	Bermasalah serius
51 - 100	E (Kurang Sekali)	Sangat bermasalah

5. Berdasarkan nilai masalah di setiap topik akan dijadikan acuan dalam memberikan layanan maupun nasihat dalam menyelesaikan permasalahan yang sedang dialami oleh peserta didik.

b) Analisis item masalah.

Analisis ini bertujuan untuk mengetahui item masalah yang pada umumnya dialami peserta didik.

Langkah-langkah analisis item masalah :

- 1) Menjumlahkan banyaknya peserta didik yang mempunyai item masalah yang sama.

$$\sum_{j=1}^k Z_j = Z_1 + Z_2 + Z_3 + \dots + Z_k$$

- 2) Mencari persentase masalah dengan mencari rasio antara banyaknya peserta didik yang bermasalah untuk item masalah tertentu dengan jumlah siswa, dirumuskan sebagai berikut :

$$X = \frac{\sum_{j=1}^k Z_j}{k} \times 100\%$$

Keterangan

X : Persentase per-topik masalah

Z_j : Banyaknya responden yang menjawab pada satu item masalah yang sama.

k : Banyak reponden.

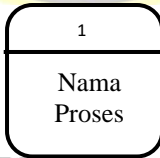
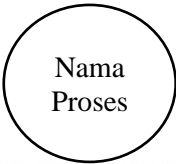
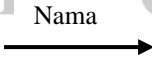
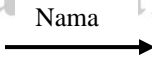

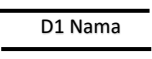
2.2.5 Diagram Alir Data (*Data Flow Diagram*)

Diagram alir data digunakan dalam pendekatan dalam menentukan, membangun dan juga menggambarkan model dari suatu sistem. Dalam pengembangan sistem dengan metode *system development life circle* (SDLC), Diagram alir data bisa digunakan dalam menentukan, membangun dan

memodelkan sistem (Ibrahim dan Yen, 2010). Dalam rekayasa software *Data flow diagram* (DFD) sangat diperlukan dalam memodelkan pemrosesan data (Zhang dkk, 2018).

Data flow diagram adalah salah satu alat pembuatan model yang digunakan untuk menggambarkan sistem sebagai suatu jaringan proses fungsional yang dihubungkan dengan alur data. Dalam merancang aplikasi, *database* sangat penting, dalam pengembangan database DFD sangat membantu dalam menentukan item data, menentukan catatan, mengidentifikasi *primary key* (Al Ashry, 2017). *Data flow diagram* digunakan dalam menentukan, membangun dan memvisualisasi model suatu sistem dengan gambar (Ibrahim dan Yen, 2011). Notasi dalam *Data flow diagram* dibagi menjadi 4 komponen (Wieringa, 2003). Simbol yang digunakan dalam diagram alir data ditunjukkan pada Tabel 2.2.

Tabel 2.2. Tabel Elemen Diagram Alir Data dalam System Analysis And Design (Dennis dkk, 2012)

Elemen <i>Data Flow Diagram</i>	Simbol Gane dan Sarson	Simbol DeMarco dan Yourdon
Proses		
Aliran data		
<i>Data store</i>		
Entitas Luar	