

BAB II

LANDASAN TEORI

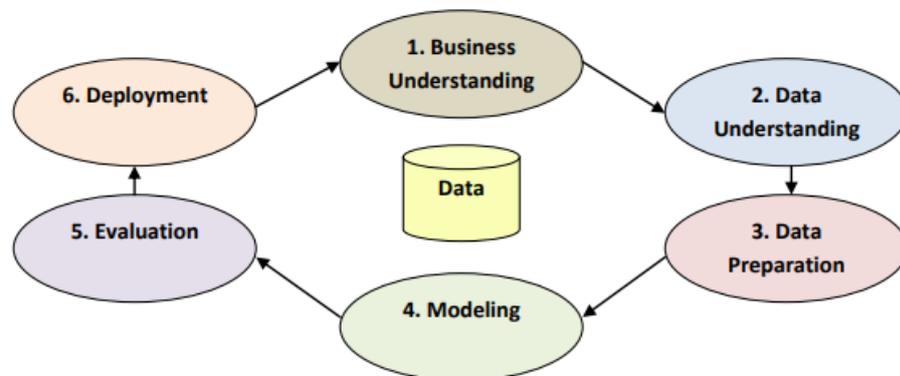
Dalam bab ini penulis akan membahas tinjauan teori yang terkait dengan penelitian yang dilakukan.

2.1 Data Mining

Pengertian data mining menurut *Gartner Group* adalah proses menemukan korelasi, pola, dan tren baru yang bermakna dengan menyaring sejumlah besar data yang disimpan dalam repositori, menggunakan teknologi pengenalan pola serta statistik dan teknik matematika (Larose, 2005). Data Mining merupakan perpaduan dari statistik terapan, logika, kecerdasan buatan, *machine learning* dan sistem manajemen data (North, 2012).

Proses data mining berdasarkan CRISP-DM (*Cross Industry Standard Process for Data Mining*) pada Gambar 1, terdiri dari 6 fase, yaitu:

- a. *Business/organizational Understanding Phase* (Fase Pemahaman Bisnis)
- b. *Data Understanding Phase* (Fase Pemahaman Data)
- c. *Data Preparation Phase* (Fase Pengolahan Data)
- d. *Modeling Phase* (Fase Pemodelan)
- e. *Evaluation Phase* (Fase Evaluasi)
- f. *Deployment Phase* (Fase Penyebaran).



Gambar 1. Fase CRISP-DM

Sumber: Data mining for the masses, North (2012)

a. *Business/Organizational Understanding Phase*

Fase *Business Understanding* adalah merupakan langkah awal melakukan CRISP DM (*Cross Industry Standard Process for Data Mining*). Pada fase ini pengetahuan akan objek bisnis atau organisasi sangat dibutuhkan, bagaimana membangun atau mendapatkan data, mencocokkan tujuan pemodelan sehingga model terbaik dapat dibangun. Kegiatan yang dilakukan antara lain: menentukan tujuan dan persyaratan dengan jelas secara keseluruhan, menerjemahkan tujuan tersebut serta menentukan pembatasan dalam perumusan masalah data mining, dan selanjutnya mempersiapkan strategi awal untuk mencapai tujuan tersebut (Larose, 2006).

b. *Data Understanding Phase* (Fase Pemahaman Data)

Pada fase ini proses mengumpulkan, mengidentifikasi, dan memahami data yang dimiliki organisasi sebelum dilanjutkan dengan tahapan selanjutnya *Data preparation*.

c. *Data Preparation Phase* (Fase Preparasi Data)

Data muncul dalam berbagai bentuk dan format antara lain: numerik, paragraph teks, Gambar yang berupa bagan, grafik dan peta, narasi seperti komentar pada survey kepuasan pelanggan atau transkrip kesaksian. Preparasi data melibatkan sejumlah kegiatan termasuk didalamnya adalah menggabungkan dua atau lebih kumpulan data, mengurangi kumpulan data menjadi hanya variabel data tertentu, membersihkan data dari anomali seperti pengamatan outlier atau data yang hilang, atau memformat ulang data untuk tujuan konsistensi.

d. *Modeling Phase* (Fase Pemodelan)

Model dalam data mining adalah representasi komputerisasi dari pengamatan dunia nyata. Model adalah algoritma untuk mencari, mengidentifikasi, dan menampilkan pola atau pesan apa pun dalam data.

e. *Evaluation Phase* (Fase Evaluasi)

Fase evaluasi digunakan untuk menentukan seberapa berharga model yang diterapkan pada tahap sebelumnya dan apa yang mungkin dilakukan dengan model tersebut.

f. *Deployment Phase* (Fase Penyebaran)

Aktivitas dalam fase ini termasuk menyiapkan otomatisasi model yang terintegrasi dengan sistem informasi manajemen atau operasional yang ada, memasukkan pembelajaran baru dari penggunaan model kembali ke dalam model untuk meningkatkan akurasi dan kinerja, dan pemantauan serta pengukuran hasil penggunaan model. Dengan mengkomunikasikan fungsi dan kegunaan model

secara jelas kepada pemangku kepentingan, secara menyeluruh menguji dan membuktikan model, kemudian merencanakan dan memantau implementasinya,

Peran utama data mining menurut (Larose, 2006) adalah:

a. Estimasi

Estimasi merupakan proses untuk menemukan suatu karakteristik penting dari data dalam suatu basis data dengan cara menduga atau menebak dari sebuah nilai yang belum diketahui. Model yang digunakan untuk estimasi di antaranya adalah: *Linier Regression, Neural Network (NN), Support Vector Machine, Generalized Liniear Model (GLM)*.

b. *Forecasting*/prediksi

Proses untuk menemukan jenis, pola atau nilai dari data dengan memperkirakan hal tersebut yang belum diketahui untuk masa mendatang. Model yang digunakan untuk *forecasting* di antaranya adalah: *Linier Regression, Neural Network (NN), Support Vector Machine, Generalized Liniear Model (GLM)*.

c. Klasifikasi

Klasifikasi merupakan proses menemukan sebuah model atau fungsi yang mendeskripsikan dan membedakan data ke dalam kelas-kelas. Klasifikasi melibatkan proses pemeriksaan karakteristik dari objek dan memasukkan objek ke dalam salah satu kelas yang sudah didefinisikan sebelumnya. Model yang digunakan untuk klasifikasi di antaranya adalah: *Decision tree, Naïve Bayes, K-Nearest Neighbour (KNN), Logistic Regresion (LoR), Linear Discriminant Analysis (LDA)*.

d. Klastering

Klastering merupakan pengelompokan data tanpa berdasarkan kelas data tertentu ke dalam kelas objek yang sama. Sebuah klaster adalah kumpulan record yang memiliki kemiripan suatu dengan yang lainnya dan memiliki ketidakmiripan dengan record dalam klaster lain. Tujuannya adalah untuk menghasilkan pengelompokan objek yang mirip satu sama lain dalam kelompok-kelompok. Semakin besar kemiripan objek dalam suatu klaster dan semakin besar perbedaan tiap klaster maka kualitas analisisnya semakin baik. Model yang digunakan untuk klastering di antaranya adalah: *K-Means*, *Kmedoids*, *Selforganizing Map (SOM)*, *Fuzzy C-means*.

e. Asosiasi

Asosiasi merupakan proses menemukan atribut yang muncul dalam suatu waktu. Dalam dunia bisnis disebut analisis keranjang belanja (*market basket analysis*). Tugas asosiasi berusaha untuk mengungkap aturan untuk mengukur hubungan antara dua atau lebih atribut. Model yang digunakan untuk asosiasi di antaranya adalah: *Coefficient of Corelation*, *Chi Square*, *FP-Growth*

2.2 Algoritma *Decision tree* dan *Naïve Bayes*

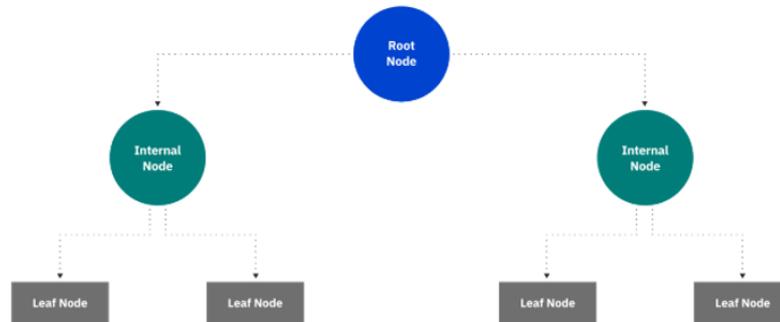
Machine learning merupakan sains (dan seni) memprogram komputer dengan belajar dari data. *Machine learning* memungkinkan sistem membuat keputusan secara mandiri tanpa dukungan dari eksternal dalam bentuk apa pun (Theobald, 2017).

Machine Learning terdiri dari *supervised learning*, *unsupervised learning* dan *reinforcement learning*. *Supervised learning* merupakan jenis yang populer

untuk melakukan operasi *machine learning* dan banyak digunakan untuk data di mana ada pemetaan yang tepat antara data input-output. Kumpulan data diberi label, artinya algoritma mengidentifikasi fitur secara eksplisit dan melakukan prediksi atau klasifikasi yang sesuai. Pada algoritma *unsupervised learning*, data tidak secara eksplisit diberi label ke dalam kelas yang berbeda (tidak memiliki label pada data). Model mampu belajar dari data dengan menemukan pola implisit. Algoritma *unsupervised learning* mengidentifikasi data berdasarkan kepadatan, struktur, segmen serupa, dan fitur serupa lainnya.

Reinforcement learning berbeda dengan *supervised* maupun *unsupervised learning*. Algoritma ini dimaksudkan untuk membuat komputer dapat belajar sendiri dari lingkungan (*environment*) melalui sebuah agen. Jadi komputer akan melakukan pencarian sendiri (*selfdiscovery*) dengan cara berinteraksi dengan lingkungan.

Decision tree adalah jenis algoritma *supervised learning* yang digunakan untuk masalah klasifikasi dan regresi. Terdapat dua tipe *decision tree* yaitu *classification tree* dan *regression tree*. Pada tipe *classification tree*, hasil yang diprediksi bersifat diskrit atau berlawanan contohnya ya atau tidak, kalah atau menang benar atau salah, sedangkan *regression tree* menghasilkan data kontinu atau masih berhubungan. Hasilnya pun bisa dianggap sebagai bilangan nyata, seperti durasi rawat inap pasien hingga harga rumah di suatu daerah. Sebuah *decision tree* memiliki hirarki seperti halnya struktur pohon, yaitu: *root node*(simpul akar), *branch* (cabang), *internal node* (simpul internal) dan *leaf node*(simpul daun) seperti dalam Gambar 2 (<https://www.ibm.com>, n.d.).



Gambar 2. Hirarki *Decision tree*

Sumber: (<https://www.ibm.com>, n.d.)

Decision tree adalah *top-down* pohon rekursif dari algoritma induksi, yang menggunakan ukuran seleksi atribut untuk memilih atribut yang diuji. Algoritma *decision tree* mencoba untuk meningkatkan akurasi dengan menghapus cabang-cabang pohon yang mencerminkan *noise* dalam data (Han *et al.*, 2011).

Beberapa kelebihan *decision tree* dalam menyusun algoritma untuk *machine learning*, yaitu:

- a. cocok untuk visualisasi data karena mudah dipahami dan diinterpretasikan,
- b. data yang digunakan dalam prosesnya hanya memerlukan persiapan minimal, tidak melalui tahapan *one-hot encoding* (kode biner 0 atau 1),
- c. struktur yang ada bisa ditambahkan opsi baru,
- d. mampu memilih opsi yang terbaik dari seluruh opsi yang tersedia,
- e. dapat bekerja dengan variabel numerik maupun kategoris.
- f. pemilihan variabel dilakukan secara otomatis. Itu berarti, jika terdapat variabel yang tidak penting, tidak akan mempengaruhi hasil akhir meski sebenarnya ada variabel yang saling berhubungan (multikolinearitas).

Kelemahan utama *decision tree* terletak pada struktur itu sendiri. Sifat struktur *decision tree* yang sangat terbuka dapat membuatnya menjadi sangat rumit,

Dalam pembuatan *decision tree* perlu diperhatikan hal-hal sebagai berikut (Santosa & Umam, 2018):

1. Variabel/atribut mana yang akan dipilih dalam pemisahan objek
2. Urutan yang akan dipilih terlebih dahulu
3. Struktur pohon
4. Kriteria pemberhentian
5. *Pruning* atau penghapusan node yang tidak berguna

Pemilihan atribut/variabel yang digunakan sebagai node/simpul adalah berdasarkan atribut yang bisa menghasilkan *decision tree* yang terkecil ukurannya atau yang bisa memisahkan objek menurut kelasnya. Dengan kata lain atribut yang dipilih adalah yang menghasilkan simpul yang paling “*purest*” (paling bersih/murni). Ukuran kemurnian yang digunakan disebut informasi dan diukur dalam satuan yang disebut bit (Witten *et al.*, 2017).

Beberapa kriteria pemilihan atribut (Santosa & Umam, 2018) yaitu :

1. *Information Gain*

Dalam memilih atribut untuk memecah objek dalam beberapa kelas, harus dipilih atribut yang menghasilkan *information gain* yang terbesar.

Sebelum menghitung *information gain* perlu dihitung terlebih dahulu nilai informasi dalam satuan bit dari suatu kumpulan objek dengan menggunakan konsep entropi. Entropi yaitu merupakan pengukuran ketidak pastian data. Entropi rendah

menunjukkan label data cukup homogen sedangkan entropi tinggi jika probabilitas sampel untuk berada ke dalam 2 (dua) kelas sebanding,

Rumus entropi:

$$\text{Entropi}(y) = -(\sum p_i \log_2 p_i) \quad (1)$$

p_i : proporsi /probabilitas kelas 1, 2....n, dalam output y

Information gain dihitung berdasarkan output data atau variabel dependen y yang dikelompokkan berdasarkan suatu atribut/variabel.A, dengan rumus sebagai berikut :

$$\text{Gain}(y,A) = \text{entropi}(y) - \sum_{i=1}^n \frac{Y_c}{y} \text{entropi}(y_c) \quad (2)$$

2. Gain Ratio

Untuk menghitung gain ratio, sebelumnya harus diketahui *split information* dengan rumus sebagai berikut:

$$\text{SI}(S, A) = - \sum_{i=1}^n \frac{S_i}{S} \log_2 \frac{S_i}{S} \quad (3)$$

SI: split information

S1 sampai Sc: pemecahan S dengan menggunakan atribut A yang memiliki nilai banyak C nilai

Rumus Gain Ratio:

$$\text{Gain ratio}(S, A) = \frac{\text{Gain}(S,A)}{\text{SI}(S,A)} \quad (4)$$

3. Gini Index (indeks gini)

Indeks Gini, di sisi lain, dikembangkan secara independen dengan tujuan awalnya adalah untuk menilai penyebaran pendapatan negara-negara tetapi kemudian disesuaikan untuk bekerja : sebagai heuristik untuk *splitting optimization*.

Rumus *Gini index* adalah:

$$IG(A) = 1 - \sum_{k=1}^c p_k^2 \quad (5)$$

k: kelas atribut/variabel

c: Jumlah kelas variabel y

P_k : Proporsi jumlah kelas dalam atribut K terhadap jumlah C dalam atribut

Setelah mengenal *information gain*, *gain ratio* dan *gini index*, dalam memilih atribut untuk memisahkan data pada *decision tree* berawal dari atribut yang memberi *information gain* terbesar, atau *gain ratio*, *gini index* yang paling kecil.

Naïve bayes adalah algoritma *machine learning* untuk melakukan klasifikasi yang berakar pada teori probabilitas dan statistik yang ditemukan oleh ilmuwan asal Inggris yaitu Thomas Bayes (Webb, 2020). Ciri khas dari *naïve bayes* adalah algoritma klasifier ini memiliki asumsi yang kuat (naif) terhadap independensi dari masing-masing kondisi/kejadian. Pada algoritma *naïve bayes*, setiap kelas keputusan akan menghitung probabilitas dengan syarat bahwa kelas keputusan tersebut benar dan juga algoritma ini mengasumsikan bahwa atribut objek adalah pelaku independen. *Naive bayes* berfungsi memprediksi probabilitas di masa depan berdasarkan pengalaman sebelumnya, sehingga dapat digunakan untuk pengambilan keputusan.

Rumus Teorema Bayes (Santosa & Umam, 2018)

$$P(h_j|x) = \frac{p(x|h_j)P(h_j)}{p(x)} \quad (6)$$

Atau

$$\text{Posterior} = \frac{\text{Prior} * \text{likelihood}}{\text{Evidence}} \quad (7)$$

x : data dengan kelas yang belum diketahui

h_j : hipotesis data h_j merupakan suatu kelas spesifik

$p(h_j|x)$: yang akan dihitung, probabilitas hipotesis sesudah observasi data (*posteriori probability*)

$P(x|h_j)$: probabilitas data x berdasar hipotesis h_j (*likelihood*)

$p(h_j)$: probabilitas hipotesis sebelum observasi data (*prior probability*)

$P(x)$: probabilitas x (*normalizing constant*)

Rumus teorema bayes jika memiliki lebih dari 1 (satu) variabel yaitu :

$$P(h_j|x_1, x_2, \dots, x_n) = \frac{\sum_{i=1}^n p(x_i|h_j)p(h_j)}{p(x_1, x_2, \dots, x_n)} \quad (8)$$

Algoritma *Naive bayes* digunakan di berbagai aspek kehidupan yaitu:

- a. *Real time prediction*
- b. *Multiclass prediction*, memprediksi probabilitas beberapa kelas variabel target
- c. *Text classification*, memiliki tingkat keberhasilan yang lebih tinggi dibandingkan dengan algoritma lain
- d. *Recommendation system*, digunakan untuk Tesis data mining untuk menyaring informasi yang tidak terlihat dan memprediksi apakah pengguna menginginkan sumber daya yang diberikan atau tidak.

Algoritma ini memiliki beberapa tipe antara lain sebagai berikut:

- a. *Bernoulli Naive bayes*

Algoritma ini digunakan ketika data sesuai dengan distribusi bernoulli multivariat, diimana prediktornya adalah variabel boolean hanya benar atau salah.

- b. *Naive bayes Multinomial*

Algoritma ini sering digunakan untuk memecahkan masalah klasifikasi dokumen. Contohnya, jika ingin menentukan apakah suatu dokumen termasuk dalam suatu kategori dan algoritma ini digunakan untuk memilahnya. *naive bayes* menggunakan frekuensi kata-kata sekarang sebagai fitur.

c. *Gaussian Naive bayes*

Digunakan jika prediktor tidak diskrit namun memiliki nilai kontinu dan prediktor tersebut diasumsikan sebagai sampel dari distribusi *gaussian*.

Kelebihan Algoritma *Naive bayes* adalah:

- a. Hanya memerlukan jumlah data sedikit yang dibutuhkan untuk klasifikasi.
- b. Cepat, efisien, dan mudah untuk dibuat.
- c. Kokoh pada atribut yang tidak relevan.

Kekurangan Algoritma *Naive bayes* adalah:

- a. Independensi antar atribut membuat akurasi menjadi berkurang.
- b. Tidak berlaku jika nilai probabilitasnya adalah nol (0).

Dari kedua model algoritma yang dijelaskan, pada saat pembahasan Tesis maka untuk pemilihan model yang terbaik yang akan diterapkan di Universitas Ma Chung dalam melakukan prediksi kelulusan tepat waktu adalah berdasarkan persentase akurasinya.

2.3 Confusion Matrix dan ROC Curve

Decision tree dan *naive bayes* merupakan algoritma klasifikasi dalam *data mining*. Pada fase evaluasi untuk algoritma klasifikasi ada 2 (dua) cara yang bisa digunakan (Field, 2017) yaitu: *Confusion Matrix* dan *Receiver Operating Characteristic (ROC) atau Precision-Recall Curve*. *Confusion matrix* digunakan

untuk membandingkan performa antara nilai prediksi dan nilai aktual. Ada 4 (empat) istilah yang merupakan representasi hasil proses klasifikasi pada *confusion matrix* yaitu *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN).

Tabel 3. *Confusion Matrix*

		Predicted Class	
		Yes (+)	No (-)
Actual	Yes (+)	True Positive (TP)	False Negative (FN)
	No (-)	False Positive (FP)	True Negative (TN)

sumber: Field (2017)

True Positive (TP): jika prediksi positif dan kenyataan positif

True Negative (TN): jika prediksi negatif dan kenyataan negatif

False Positive (FP): jika prediksi positif tetapi kenyataan negatif

False Negative (FN): jika prediksi negatif dan kenyataan juga negatif

Ada 4 (empat) pengukuran yang dapat dilakukan dengan *confusion matrix* yaitu: akurasi, presisi, *recall/sensitivity* dan *F1 score*. Pengukuran akurasi adalah seberapa akurat model dalam mengklasifikasi secara benar. Presisi adalah akurasi antara data yang diminta dengan prediksi. *Recall* menggambarkan keberhasilan model dalam menemukan kembali sebuah informasi. *F1 score* menggambarkan perbandingan rata-rata presisi dan *recall* yang dibobotkan, akurasi digunakan sebagai acuan performansi algoritma jika jumlah data FN dan FP mendekati (simetris), jika tidak maka *F1 score* yang digunakan.

$$\text{Rumus akurasi: } \frac{TP+TN}{TP+TN+FP+FN} \quad (9)$$

$$\text{Rumus presisi: } \frac{TP}{TP+FP} \quad (10)$$

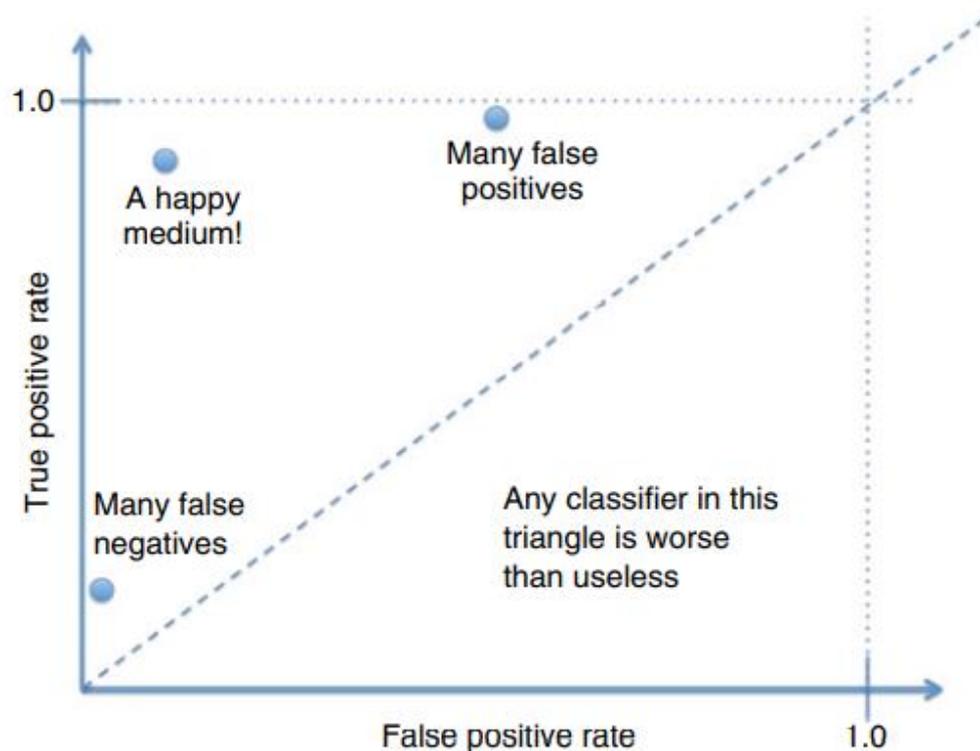
$$\text{Rumus recall: } TP/TP + FN \quad (11)$$

$$\text{Rumus F1 score: } (2 \times \text{recall} \times \text{presisi}) / (\text{recall} + \text{presisi}) \quad (12)$$

Receiver Operating Characteristic (ROC) atau Precision-Recall Curve dapat menggambarkan kinerja pengklasifikasi tanpa memperhatikan distribusi kelas (Witten *et al.*, 2017). *ROC curve* dibuat berdasarkan nilai yang telah didapat pada perhitungan dengan *confusion matrix*, yaitu antara *False Positive Rate (FPR)* dengan *True Positive Rate (TPR)*.

$$\text{Rumus False Positive Rate (FPR) = } FP / (FP + TN) \quad (13)$$

$$\text{Rumus True Positive Rate (TPR) = } TP / (TP + FN) \quad (14)$$



Gambar 3. ROC curve

Sumber: Field (2017)

Kurva *Receiving Operator Characteristic* (ROC) yang dihasilkan mendekati garis baseline ($x=y$) atau garis yang melintang dari titik 0,0 maka dikatakan jelek, dikatakan bagus jika mendekati titik 0,1 (Gambar 3).

Nilai yang dihasilkan dapat dikategorikan sebagai nilai dalam *Area Under Curve* (AUC) (Gorunescu, 2011) adalah sebagai berikut:

0,9 – 1,00: *excellent*

0,8 – 0,9: *good*

0,7 – 0,8: *fair*

0,6 – 0,7: *poor*

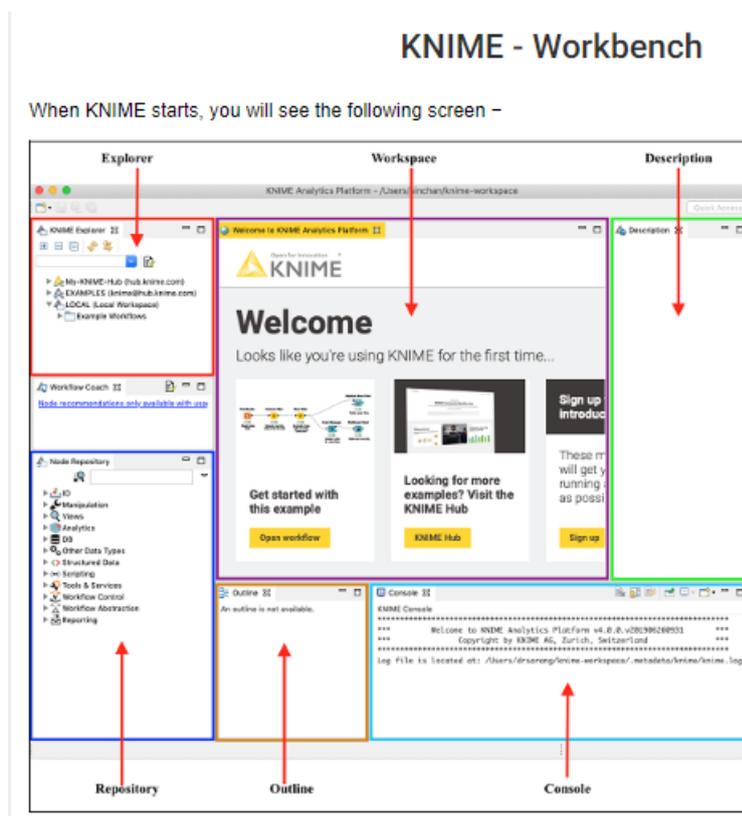
0,5 – 0,6: *failure*

2.4 Knime

Konstanz Information Miner (KNIME) adalah platform *data mining* untuk analisis, pelaporan dan integrasi data yang termasuk software *opensource* (<https://www.knime.com/software-overview>, n.d.). *Knime* mulai dikembangkan tahun 2004 oleh tim pengembang perangkat lunak dari Universitas Konstanz, yang dipimpin oleh Michael Berthold, yang awalnya digunakan untuk penelitian di industri farmasi. Platform ini banyak digunakan sejak tahun 2006, sehingga tahun 2017 masuk ke *magic quadrant for data science platform* (Gartner Group).

Penulis akan menjelaskan tentang tampilan dan proses analisis menggunakan aplikasi *Knime* (www.tutorialspoint.com/knime, n.d.) secara garis besar sebagai berikut :

1. ketika menjalankan knime akan tampil di layar *Knime-Workbench*, yang berisikan *Workspace*, *outline*, *repositori node*, *explorer*, *console* dan *description* seperti pada Gambar 4.



Gambar 4. *Workbench Knime*

Sumber: ([Www.Tutorialspoint.Com/Knime](http://www.Tutorialspoint.Com/Knime), n.d.)

Workspace merupakan kertas kerja dalam melakukan analisis, *outline* merupakan tampilan miniatur kerangka kerja, *Node* mendefinisikan jenis fungsionalitas tertentu yang digunakan dalam workflow, konsol menyediakan tampilan berbagai pesan saat menjalankan alur kerja, *explorer* mencantumkan ruangan dimana file *Knime* berada dan *description* menjelaskan item yang sedang dipilih di *workspace*.

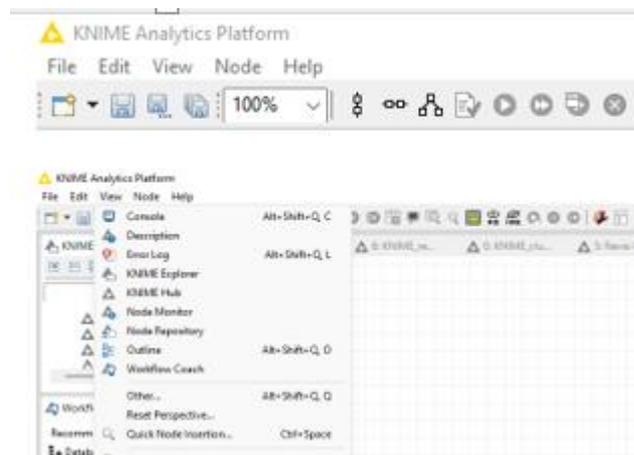
2. Toolbar yang berisikan ikon untuk memfasilitasi tindakan cepat dengan melakukan klik mouse seperti tampak dalam Gambar 5.



Gambar 5. *Toolbar Knime*

Sumber: ([Www.Tutorialspoint.Com/Knime](http://www.tutorialspoint.com/knime), n.d.)

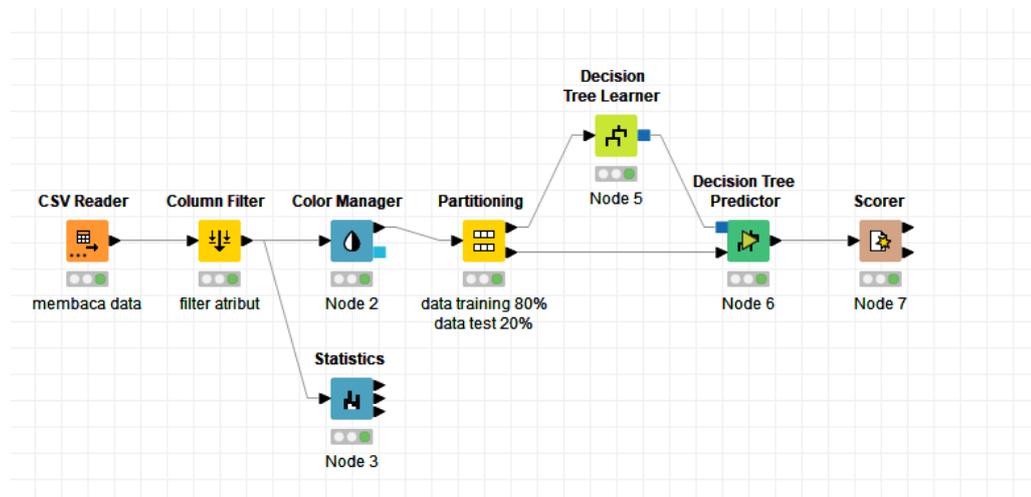
3. Menu yang terdapat pada *Knime* terdiri dari *File*, *Edit*, *View*, *Node* dan *Help* seperti tampak pada Gambar 6.



Gambar 6. *Menu Knime*

Sumber: ([Www.Tutorialspoint.Com/Knime](http://www.tutorialspoint.com/knime), n.d.)

4. Untuk melakukan analisis dengan *Knime* sangatlah mudah, cukup menggunakan *Node Analytics* yang tersedia, contohnya menentukan berbagai algoritma *machine learning* seperti *decision tree*, *naïve bayes*, *clustering*, *ensemble learning* dan sebagainya.



Gambar 7. Node Knime

Sumber: diolah Penulis (2022)

2.5 Inovasi

Statistik Direktorat Pendidikan Tinggi (Kemendikbud, 2020) menyebutkan di tahun 2020 berdasarkan bentuk, Indonesia memiliki 4593 perguruan tinggi yang terdiri dari 3403 pendidikan tinggi akademik (dikti) dan 1190 pendidikan vokasi (diksi), dari jumlah tersebut perguruan tinggi swasta (PTS) sebanyak 3044 (66,27%) dan sisanya 1.549 (33,72%) adalah perguruan tinggi negeri, perguruan tinggi keagamaan dan perguruan tinggi kedinasan. Total perguruan tinggi yang sudah terakreditasi sebanyak 2713 yaitu 59,01%. Diketahui pula bahwa jumlah program studi sebanyak 29.413 dan yang sudah terakreditasi sebanyak 20.691 atau 70,3%. Adapun jumlah mahasiswa terdaftar sebanyak 8.483.213 orang dan jumlah lulusan sebanyak 1.535.074 orang. Dari data Kemendikbud tersebut tampak bahwa persaingan antar perguruan tinggi di Indonesia sangatlah ketat.

Universitas Ma Chung sebagai penyelenggara pendidikan tinggi haruslah memiliki strategi untuk menciptakan keunggulan bersaing. Strategi menurut Learned

(Rangkuti, 2013) merupakan alat untuk menciptakan keunggulan bersaing. Salah satu strategi yang dilakukan untuk meningkatkan keunggulan kompetitif adalah dengan melakukan inovasi. Inovasi memiliki fungsi yang khas bagi organisasi, dengan inovasi menciptakan baik sumber daya produksi baru maupun pengolahan sumber daya yang ada dengan peningkatan nilai potensi untuk menciptakan modal (Drucker, 2012).

Menurut (Keeley, 2013) terdapat sepuluh tipe inovasi seperti dalam Gambar 8, yaitu:

a. *Profit Model*

Inovasi ini menjelaskan cara baru perusahaan menghasilkan uang, dengan memanfaatkan penawaran perusahaan dan sumber lainnya. Dalam hal ini pemahaman secara mendalam perusahaan terhadap apa yang dilakukan pelanggan akan menimbulkan peluang mendapatkan uang.

b. *Network*

Inovasi ini menjelaskan cara perusahaan memanfaatkan networknya untuk bersama-sama menghasilkan value. Perusahaan akan berkolaborasi dengan perusahaan lain.

c. *Structure*

Perusahaan berinovasi dengan menggunakan semua aset baik yang berwujud ataupun tidak untuk menghasilkan nilai ataupun bisnis yang optimal

d. *Process*,

Inovasi yang dilakukan berbasis proses bisnis, mencakup metodologi ataupun kemampuan yang lebih tinggi daripada perusahaan yang sejenis. Perusahaan memiliki keunikan dalam menyampaikan produk ataupun memberikan layanan.

e. Product Performance

Inovasi yang dilakukan dengan mengembangkan fitur dan fungsi yang berbeda dari perusahaan lain. Inovasi kinerja produk membahas nilai, fitur, dan kualitas barang yang ditawarkan.

f. Product System

Inovasi dalam sistem penciptaan produk, perusahaan membuat pelengkap dari produk dan layanan. Perusahaan berinovasi membuat bundling antara produk dan layanan, untuk menciptakan sistem yang kuat dan terukur. Bundling produk, atau mengambil beberapa produk terkait dan menjualnya dalam satu paket, adalah contoh inovasi Sistem Produk

g. Service

Inovasi berbasis layanan, mendukung dan memperkuat nilai produk.

Inovasi layanan memastikan dan meningkatkan utilitas, kinerja, dan nilai dari suatu produk. Produk lebih mudah untuk dicoba, digunakan, dan dinikmati sehingga membuat pelanggan datang kembali.

h. Channel

Inovasi saluran distribusi mencakup semua cara menghubungkan produk perusahaan dengan pelanggan. Saluran *e-commerce* telah muncul sebagai kekuatan dominan tetapi saluran tradisional seperti toko masih penting. Inovasi saluran untuk memastikan bahwa pengguna dapat membeli apa yang mereka

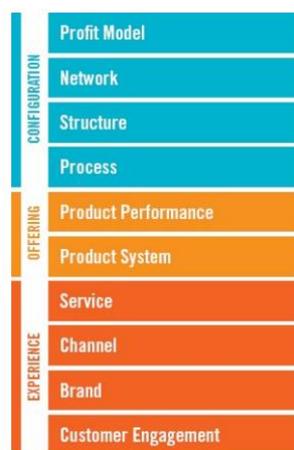
inginkan, kapan dan bagaimana mereka menginginkannya, dengan sedikit biaya dan tetap mendapat kepuasan yang maksimal.

i. Brand

Inovasi merek membantu memastikan bahwa pelanggan dan pengguna mengenali, mengingat, dan memilih produk perusahaan daripada pesaing.

j. Customer Engagement

Keakraban dengan pelanggan juga dapat di inovasikan sehingga menghasilkan sebuah *value* yang sangat dihargai oleh pelanggan.



Gambar 8. Ten Types of Innovation

Sumber: (Keeley, 2013)

Penelitian yang penulis lakukan di Bagian Administrasi Akademik (BAA) dengan melakukan prediksi kelulusan tepat waktu mahasiswa Universitas Ma Chung dan memberikan rekomendasi terkait “*early warning system*” ataupun notifikasi di *Ma Chung Academic Integrated System (MACIS)* yang berguna bagi mahasiswa, dosen penasihat akademik dan institusi adalah merupakan inovasi proses bagi Universitas Ma Chung. Dengan sistem tersebut diharapkan akan berguna dalam proses

pembelajaran, memberikan peringatan bagi mahasiswa terkait kelulusan tepat waktu, bagi dosen penasihat akademik dalam menjalankan fungsinya untuk mengarahkan mahasiswa dalam bidang akademik dan bagi institusi agar dapat membuat kebijakan yang mendorong mahasiswa lulus tepat waktu. Peringatan bagi mahasiswa dan peran dosen penasihat akademik erat kaitannya dengan inovasi *service* dan *customer engagement*.

2.6 Efektivitas dan Produktivitas Pembelajaran

Efektivitas adalah melakukan pekerjaan yang benar (*doing the right things*), sedang efisien adalah melakukan pekerjaan dengan benar (*doing things right*) (Drucker, 2017). Dalam standar nasional pendidikan tinggi, pengertian efektif dalam standar proses adalah capaian pembelajaran lulusan diraih secara berhasil guna dengan mementingkan internalisasi materi secara baik dan benar dalam kurun waktu yang optimum, sedangkan pembelajaran adalah proses interaksi mahasiswa dengan Dosen dan sumber belajar pada suatu lingkungan belajar (Mendikbud RI, 2020). Dengan demikian efektivitas pembelajaran dapat diartikan sebagai ukuran keberhasilan dari proses interaksi mahasiswa dan dosen.

Efektivitas pembelajaran diukur dengan indikator (Slavin, 2018) sebagai berikut:

1. Kualitas pembelajaran (*quality of assurance*), yaitu seberapa besar kadar informasi yang disajikan sehingga mahasiswa dengan mudah dapat mempelajarinya atau tingkat kesalahannya semakin kecil. Semakin kecil tingkat kesalahan yang dilakukan berarti semakin efektif pembelajaran.

2. Kesesuaian tingkat pembelajaran (*appropriate level of instruction*) yaitu sejauh mana dosen memastikan tingkat kesiapan mahasiswa dalam menerima materi baru.
3. Intensif yaitu seberapa besar usaha dosen memotivasi mahasiswa untuk menyelesaikan atau mengerjakan tugas-tugas dan mempelajari materi yang diberikan.
4. Waktu, yaitu waktu yg dibutuhkan untuk menyelesaikan kegiatan pembelajaran. Pembelajaran akan efektif apabila mahasiswa dapat menyelesaikan pembelajaran sesuai dengan waktu yang ditentukan.

Pengertian Produktivitas menurut Peter F. Drucker adanya keseimbangan antara semua faktor produksi sehingga input terkecil bisa memberikan output terbesar (Drucker, 2008), sedangkan menurut Handoko adalah hubungan antara masukan-masukan dan keluaran-keluaran suatu sistem produktif. Dalam teori, sering mudah untuk mengukur hubungan ini sebagai rasio keluaran dibagi masukan. Bila lebih banyak keluaran diproduksi dengan jumlah masukan sama, produktivitas naik. Begitu juga, bila lebih sedikit masukan digunakan untuk sejumlah keluaran sama, produktivitas juga naik (Handoko, 2013). Apabila dihubungkan dengan pembelajaran maka produktivitas merupakan perbandingan luaran dan masukan dari proses interaksi mahasiswa dan dosen. Dalam proses pembelajaran di perguruan tinggi maka masukan dapat berupa jumlah mahasiswa baru dan keluarannya adalah jumlah mahasiswa yang berhasil lulus.

Berdasarkan instrumen akreditasi program studi, kriteria 9. luaran dan capaian tridharma khususnya pendidikan, difokuskan pada pencapaian dan

kompetensi lulusan berupa gambaran yang jelas tentang profil dan capaian pembelajaran lulusan/kompetensi yang ditetapkan oleh program studi dan perguruan tinggi dengan mengacu KKNI. Adapun elemen dan deskripsi penilaiannya meliputi:

1. analisis pemenuhan capaian pembelajaran lulusan (CPL) yang diukur dengan metoda yang sah dan relevan, mencakup aspek: keserbacakupan, kedalaman, kebermanfaatan,
2. indeks prestasi kelulusan (IPK) lulusan,
3. prestasi mahasiswa di bidang akademik,
4. prestasi mahasiswa di bidang non akademik,
5. masa studi,
6. kelulusan tepat waktu,
7. keberhasilan studi,
8. pelaksanaan tracer studi,
9. waktu tunggu,
10. kesesuaian bidang kerja,
11. tingkat kepuasan pengguna lulusan

Dari elemen penilaian tersebut untuk menilai efektivitas dan produktivitas pembelajaran diukur berdasarkan lamanya masa studi dan persentase kelulusan mahasiswa tepat waktu.