

BAB II

KAJIAN PUSTAKA

Pada bab ini akan diuraikan tinjauan pustaka yang berkaitan dengan penelitian terbaru tentang prediksi hasil dari operasi THA, orisinalitas penelitian disertasi, serta landasan teori yang berkaitan dengan THA, pengukuran hasil dan skor fungsi panggul, serta konsep XAI.

2.1 Tinjauan Pustaka

Sejumlah studi telah dilakukan untuk memprediksi hasil dari operasi THA menggunakan pendekatan *machine learning* yang dilengkapi dengan penjelasan terhadap hasil prediksi yang dilakukan. Penelitian oleh Kuo (Kuo et al., 2022) memprediksi infeksi sendi prostetik panggul dan lutut berdasarkan data rekam medis dengan mengimplementasikan 2 (dua) tingkat arsitektur *machine learning*. *Naïve-Bayes*, *eXtreme Gradient Boosting*, *Linear Regression*, dan *Random Forest* adalah algoritma *machine learning* yang digunakan sebagai *base classifier* pada tingkat 1 sedangkan *Support Vector Machine* sebagai *meta-classifier* yang bertindak sebagai pengambil keputusan final pada tingkat 2 berdasarkan hasil prediksi di tingkat 1. Variabel pasien yang digunakan sebagai input adalah demografi, biomedik, komorbiditas, dan variabel yang berkaitan dengan proses operasi serta memiliki kinerja yang sangat baik dengan nilai AUC sebesar 0,988. Sebuah pembangkit penjelasan dibuat dengan mengekstraksi *decision path* yang konsisten dengan hasil prediksi. *Decision path* yang memiliki fidelitas tinggi dikonversi menjadi aturan *if-then* dan dibuat menjadi sebuah diagram keputusan untuk menjelaskan hasil prediksi.

Penelitian yang dilakukan oleh Klemmt (Klemmt et al., 2022) menggunakan variabel pasien yang serupa dengan penelitian Kuo untuk memprediksi revisi dari operasi revisi THA, namun pada tahun berikutnya menggunakan tambahan variabel PROM pada fase pra-operatif dan memprediksi PROM pada fase pasca-operatif

(Klemmt et al., 2023). Algoritma *machine learning* Jaringan Saraf Tiruan, *Elastic-net penalized logistic regression*, dan *Random Forest* digunakan pada kedua penelitian tersebut dengan Jaringan Saraf Tiruan memiliki AUC tertinggi masing-masing sebesar 0,85 dan 0,89. Sebagai teknik XAI-nya, keduanya menghasilkan plot dari *Relative Variable Importance* untuk mengidentifikasi parameter yang paling penting bagi model *machine learning* dalam melakukan prediksi.

Peneliti KN Kunze i(Kunze et al., 2020) melakukan prediksi perbedaan penting klinis yang minimal (*Minimal Clinical Important Difference* – MCID) pada laporan kesehatan yang dilaporkan oleh pasien dalam rentang sebelum dan sesudah menjalani operasi THA untuk mengukur kesuksesan hasil operasi. Algoritma *stochastic gradient boosting*, *Random Forest*, *Support Vector Machine*, Jaringan Saraf Tiruan, dan *Elastic-net penalized logistic regression* digunakan secara terpisah terhadap basis data dan *Random Forest* memiliki AUC tertinggi sebesar 0,97. Teknik XAI yang digunakan adalah plot *Relative Variable Importance* untuk menentukan variabel yang paling penting yang digunakan oleh algoritma dalam melakukan prediksi.

Prediksi komplikasi dan durasi operasi yang tidak lazim dilakukan oleh Lazic (Lazic et al., 2022) menggunakan algoritma *eXtreme Gradient Boosting* dengan menggunakan 12 variabel terkait demografi pasien, jenis implan, dan pengalaman dokter bedah. Nilai AUC untuk prediksi komplikasi sebesar 0,64 dan prediksi durasi operasi sebesar 0,89. Untuk penjelasan hasil prediksi model juga menggunakan teknik *feature importance* yang memberikan skor relatif terhadap parameter input berdasarkan seberapa besar kontribusi mereka dalam memprediksi sebuah target, namun tidak menunjukkan hubungan sebab-akibat antara input dan output.

Penelitian oleh Nham (Nham et al., 2023) menggunakan 10 macam algoritma yang berbasis *tree*, jaringan saraf tiruan, pendekatan statistik dan probabilitas, dan *Support Vector Machine* untuk memprediksi mortalitas setelah operasi, disposisi setelah keluar dari rumah sakit, dan lama rawat inap. Variabel yang digunakan adalah variabel spesifik pasien (seperti demografi, diagnosis, dan risiko mortalitas) dan variabel situasional (seperti lokasi operasi, jenis rumah sakit, dan jumlah

tempat tidur di rumah sakit). Untuk tiap output yang diprediksi terdapat 3 model terbaik yang berbeda, namun *Linear Support Vector Machine* dan *Decision List* adalah algoritma yang paling sering masuk dalam 3 model terbaik untuk semua output dan model yang dilatih pada variabel spesifik pasien memberikan kinerja yang lebih baik daripada yang dilatih pada variabel situasional. Tidak ada informasi terkait teknik XAI yang digunakan, namun sebagian besar algoritma yang digunakan termasuk dalam kategori *interpretable* yang sudah cukup transparan.

Dalam penelitian yang dilakukan oleh Pakarinen (Pakarinen et al., 2022), digunakan 37 variabel dari 172 variabel (dipilih manual oleh tim pakar peneliti), yang terdiri dari faktor terkait pasien dan faktor terkait operasi atau implan yang digunakan, untuk membuat 2 model untuk prediksi operasi revisi karena adanya dislokasi dalam kurun waktu 2 tahun setelah operasi THA primer. Model pertama menggunakan algoritma *Logistic Regression* yang menggunakan 14 variabel input (diseleksi berdasarkan data dari literatur) memiliki *Concordance Index* sebesar 0,67 dan R^2 sebesar 0,04 yang mengindikasikan rendahnya kemampuan prediksi. Model kedua menggunakan *Elastic-net regularization* dan memiliki nilai AUC sebesar 0,73 serta R^2 sebesar 0,02 serta memilih 9 variabel pada model final, termasuk di dalamnya variabel terkait diagnosis pra-operatif, jenis fiksasi femur, dan demografi pasien. Kedua model merupakan *interpretable model* yang secara intrinsik memiliki sifat yang transparan, di mana variabel yang terpilih oleh model dan koefisien regresi menunjukkan variabel yang berkontribusi dalam hasil prediksi sehingga tidak memerlukan penjelasan tambahan.

Hernigou (Hernigou et al., 2023) menggunakan pendekatan yang berbeda dalam memprediksi risiko dislokasi berdasarkan tipe implan dengan memanfaatkan artikel yang terkait dengan dislokasi sendi panggul yang dipublikasikan di basisdata PubMed. Ia melakukan *systematic review* terhadap artikel terkait dislokasi sendi panggul dan mendapatkan 75 artikel yang memenuhi kriteria. Dalam 75 artikel tersebut terdapat 895 ribu data operasi penggantian sendi panggul dengan sekitar 1 juta implan untuk dilakukan ekstraksi data demografi pasien, komorbid, karakteristik bedah, dan dislokasi sebanyak 26.488 kasus. Algoritma Jaringan Saraf Tiruan digunakan untuk memprediksi kemungkinan dislokasi dari empat tipe

implan (*dual mobility*, *hemiarthroplasty*, *constrained liner*, dan implan konvensional THA), dengan faktor terkait pasien, posisi implan, dan faktor terkait bedah berturut-turut sebagai variabel pada lapis tersembunyi kesatu, kedua, dan ketiga untuk menghasilkan dua macam output, yakni Dislokasi atau Tidak Dislokasi. Akurasi sistem sebesar 95% dan untuk teknik XAI yang digunakan adalah Shapley Additive Explanation (SHAP) (Lundberg & Lee, 2017) yang mempertimbangkan *outcome* dari kombinasi beberapa faktor untuk membuat peringkat tingkat kepentingan sebuah faktor terhadap hasil prediksi. Pada penggunaan implan konvensional THA terdapat 5 (lima) faktor yang memiliki skor SHAP tertinggi, yang menyatakan faktor yang paling dominan, yakni *Body Mass Index* $> 32 \text{ kg/m}^2$, riwayat penyakit neurologis, riwayat menurunnya kemampuan kognitif, ukuran diameter kepala implan 22 mm, dan posisi implan *femoral retroversion*.

Shah dan tim (Shah et al., 2021) melakukan prediksi komplikasi setelah operasi THA dengan menggunakan *tool* AutoPrognosis (Alaa & van der Schaar, 2018) yang berupa *Automated Machine Learning* dengan kemampuan untuk melakukan ansambel dari sejumlah algoritma *machine learning* yang dipilih secara otomatis untuk digunakan dan yang paling sesuai dengan karakteristik basisdata yang diinputkan. AutoPrognosis menggunakan *Bayesian Optimization* untuk memilih *hyperparameter* terbaik dari masing-masing algoritma untuk menghasilkan 10 (sepuluh) *pipeline machine learning* terbaik beserta setting *hyperparameter*-nya. Sistem ini memiliki nilai AUROC 0,73, yang diklaim paling tinggi dari lima algoritma pembandingan, yakni *Logistic Regression*, *XGBoost*, *Gradient Boosting*, *AdaBoost*, dan *Random Forest*, pada basisdata yang sama. Teknik XAI yang digunakan di sistem ini adalah *feature importance*, yang menentukan 5 (lima) faktor paling berpengaruh terhadap hasil prediksi, yakni malnutrisi, demencia, riwayat penyakit kanker, penyakit paru-paru kronis, dan cakupan asuransi kesehatan yang digunakan oleh pasien.

Terkait dengan penggunaan teknik XAI pada domain kesehatan yang lebih luas, penjelasan menggunakan *feature* atau *variable importance* juga masih menjadi pilihan dari banyak penelitian yang menghasilkan model *machine*

learning. Sebuah penelitian (Huang et al., 2024) untuk memberikan rekomendasi obat berdasarkan catatan pengobatan sebelumnya pada pasien penyakit kronis dilakukan dengan menggunakan *Graph Neural Network-based Collaborative Filtering* (GCF). Pada penelitian ini dibangun graf bipartit yang menggambarkan interaksi antara pasien dengan *item* (seperti pembelian obat, rating obat, dan lain-lain) dengan dua macam graf, dengan mekanisme *attention* dan tanpa mekanisme *attention* terhadap node tetangga. Sumber data berasal dari sebuah rumah sakit di China dan juga menggunakan basisdata publik. Presisi tertinggi sebesar 0,85 pada data rumah sakit dan 0,88 pada basisdata publik. Teknik XAI yang digunakan adalah SHAP dan LIME.

Penelitian yang dilakukan oleh Alsinglawi (Alsinglawi et al., 2022) membangun sebuah kerangka *machine learning* untuk memprediksi lama rawat inap (LoS) dari pasien kanker paru-paru yang masuk di ruang ICU berdasarkan data riwayat medis dari basisdata publik MIMIC-III. Penelitian ini menggunakan 6 (enam) algoritma penyeimbang kelas yang bertipe *oversampling* dan *undersampling* untuk kemudian dilakukan klasifikasi dengan algoritma *Random Forest*. Basisdata memiliki total 75 atribut yang kemudian dilakukan pemilihan atribut dengan metode *Recursive Feature Elimination* (RFE) untuk menghasilkan 60 atribut final. Skor AUC tertinggi sebesar 100% diraih oleh algoritma penyeimbang ADASYN dan posisi berikutnya adalah SMOTE dan SMOTETomek dengan skor 98%. Teknik XAI yang digunakan adalah SHAP dengan top-5 (lima) fitur adalah *temperature_F*, *platelets*, *ADM_Emergency*, *glucose*, dan *chloride*.

Sebuah sistem dengan teknik *deep learning* (Wani et al., 2024) juga diusulkan untuk mendeteksi penyakit kanker paru-paru dengan basisdata publik *Survey Lung Cancer* yang memiliki 16 atribut, di antaranya adalah *smoking*, *anxiety*, *shortness of breath*, *chest pain*, dan *coughing*. Sistem ini membangun sebuah *Convolutional Neural Network* (CNN) yang melakukan ekstraksi ciri terhadap data input dan melakukan prediksi dengan algoritma *XGBoost* dengan nilai F1-score 98,08. Penjelasan terhadap hasil deteksi dilakukan dengan teknik SHAP untuk menghasilkan sejumlah atribut yang paling berkontribusi terhadap hasil deteksi, di antaranya adalah *chronic disease*, *fatigue*, *peer_pressure*, *alcohol consuming*, dan

allergy.

Pada sebuah sistem klasifikasi data *arrythmia* berdasarkan data *Electrocardiogram* (ECG) (Raza et al., 2022) digunakan sebuah autoencoder, algoritma klasifikasi CNN, dan modul XAI *Gradient-weighted Class Activation Mapping* (Grad-CAM) dalam lingkungan *federated learning*. Autoencoder digunakan untuk menghilangkan derau dari sinyal input ECG sekaligus untuk mengurangi risiko *overfitting*. Algoritma klasifikasi CNN dengan 4 *convolutional layers*, 3 *max pooling layers*, 2 *fully connected layers*, dan 1 *softmax layer* untuk mengklasifikasi input ke dalam 5 (lima) kelas. Modul XAI Grad-CAM yang digunakan merupakan penjelasan *model-specific* untuk model CNN yang dimodifikasi untuk data *time-series*. Bentuk penjelasannya berupa visualisasi *heatmap* yang menunjukkan segmen pada data input yang dianggap penting oleh algoritma dalam melakukan klasifikasi.

Bidang kajian pemrosesan bahasa natural juga telah menggunakan fitur XAI untuk memfasilitasi penemukembalian data oleh pengguna non-teknis menggunakan graf pengetahuan. *Explicable Knowledge Graph* (X-KG) (Shaikh et al., 2024) merupakan pendekatan XAI yang dirancang untuk menterjemahkan bahasa *query* natural ke bahasa *query* SPARQL untuk *query* RDF (*Resource Description Framework*) yang merupakan basis dari graf pengetahuan. Basisdata yang digunakan adalah basisdata COVID-19 dan Stock Market Data serta memiliki akurasi sebesar 91,91% untuk *query* SELECT pada data COVID-19.

Pada penelitian yang lain, penggunaan model yang *interpretable*, yakni *Explainable Boosting Machine*, untuk memprediksi dan mengidentifikasi berbagai faktor risiko kehamilan dan kelahiran terbukti dapat mengungguli sejumlah algoritma *machine learning black box* lain seperti XGBoost, Random Forest, dan DNN dengan nilai AUROC rata-rata sebesar 0,73 (Bosschieter et al., 2023). Model dapat menentukan faktor yang paling menentukan untuk masing-masing empat risiko yang diidentifikasi, yakni *birthweight* untuk risiko *Shoulder dystocia*, *maternal BMI* untuk risiko *Preterm preeclampsia*, *nulliparity* untuk risiko *Severe Maternal Morbidity*, dan *race* untuk risiko *Antepartum stillbirth*.

Studi literatur tentang aplikasi AI/ML untuk prediksi hasil pasca-operasi

THA yang telah dilakukan penulis sebelumnya memperoleh hasil teknik dan pendekatan *machine learning* yang banyak digunakan untuk prediksi adalah *deep learning*, *Random Forest*, *Convolutional Neural Network*, dan *Logistic Regression* (Purbasari et al., 2023). Temuan ini juga memperkuat hasil studi lain tentang aplikasi AI/ML pada THA dan TKA dengan algoritma/pendekatan berbasis *Neural Network*, *Boosting*, *Decision Tree*, *Regression*, *Support Vector Machine*, dan *Cluster Analysis* diidentifikasi sebagai pendekatan yang dipilih untuk melakukan prediksi hasil pasca-operasi (Lopez et al., 2021).

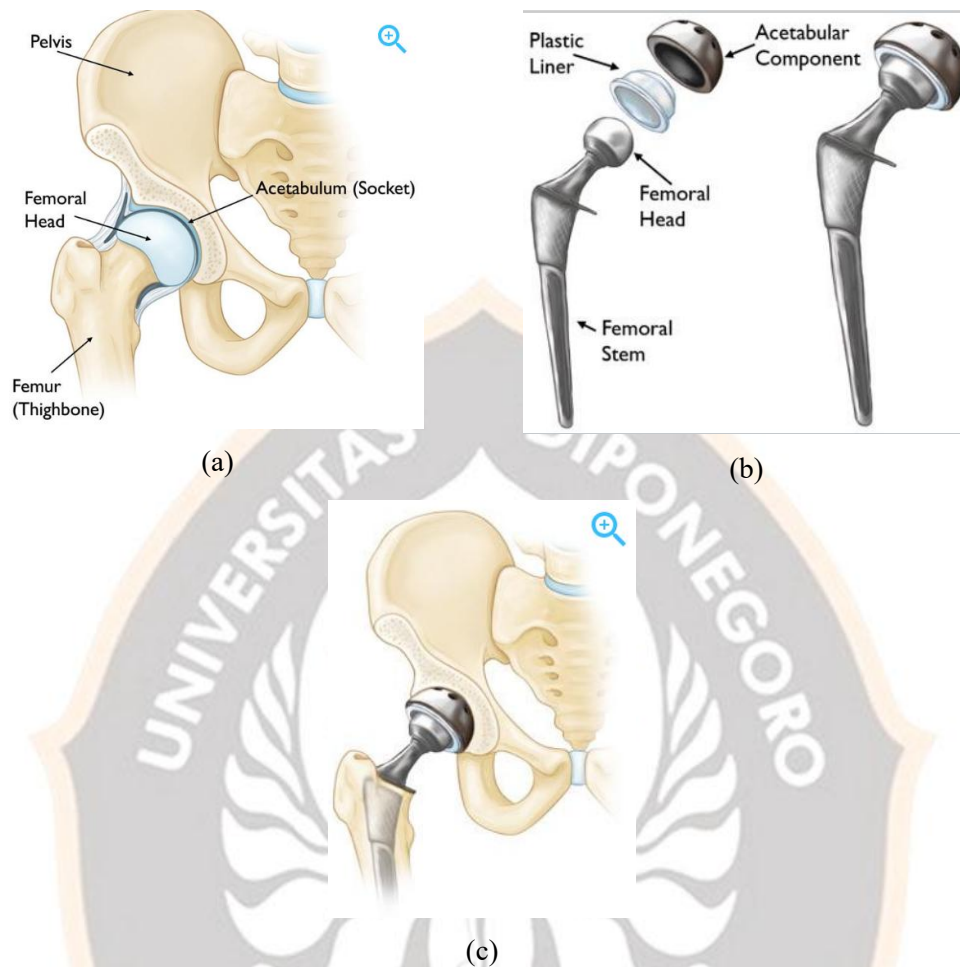
2.2 Landasan Teori

Pada bagian ini akan dijelaskan secara singkat istilah dan teori yang digunakan dalam pelaksanaan penelitian disertasi ini.

2.2.1 THA (Operasi penggantian sendi panggul total)

1. THA dan prosedurnya

Operasi penggantian sendi panggul dapat berupa penggantian sendi sebagian (hanya mengganti bagian kepala tulang paha) atau total (mengganti bagian kepala tulang paha dan acetabulum) (Parker & Pervez, 2002). Anatomi panggul pada Gambar 2.1a terdiri atas bagian pelvis dan sendi panggul yang terdiri dari bagian acetabulum dan *femoral head* di bagian pangkal *femur*. THA (atau sering juga disebut THR – *Total Hip Replacement*) adalah prosedur operasi penggantian tulang dan kartilago di bagian sendi panggul dengan komponen prostetik dengan mengganti kepala tulang paha dan permukaan sendi dari asetabulum dipasangkan pada permukaan sendi sintesis yang berbentuk mangkok. Gambar 2.1b menjelaskan komponen pada prostesis panggul dan 2.1c menggambarkan bagaimana prosthesis tersebut menggantikan sendi panggul.



Gambar 2.1. Anatomi panggul manusia (a) Anatomi panggul normal, (b) Prostesis/implan, dan (c) anatomi panggul dengan prostesis. (Sumber: <https://orthoinfo.aaos.org/en/treatment/total-hip-replacement/>)

2. Bahan Material Implan

Terdapat 5 (lima) kombinasi bahan material yang digunakan pada implan panggul, yakni (US Food and Drug Administration, 2019):

- Metal-on-Polyethylene* (MoP), di mana bagian bola (kepala tulang paha atau *femoral head*) terbuat dari metal dan bagian rongga (*liner*) terbuat dari plastik polietilen (*polyethylene*)
- Ceramic-on-Polyethylene* (CoP), di mana bagian bola terbuat dari keramik dan bagian rongga dari plastik polietilen
- Ceramic-on-Ceramic* (CoC), di mana bagian bola dan bagian rongga seluruhnya terbuat dari keramik
- Ceramic-on-Metal* (CoM), di mana bagian bola terbuat dari keramik dan

bagian rongga dari metal

- e. *Metal-on-Metal* (MoM), di mana bagian bola dan bagian rongga seluruhnya terbuat dari metal. Terhitung sejak 16 Mei 2016, negara Amerika Serikat telah meniadakan tipe ini dari daftar alat kesehatan yang disetujui oleh *Food and Drug Administration* (FDA) karena terjadi proses aus dan korosi yang menghasilkan ion metal dan memasuki aliran darah sehingga menyebabkan risiko reaksi yang tidak dapat diprediksi.

3. Perawatan dan tindak lanjut pascaoperasi THA

Kesuksesan hasil operasi juga sangat ditunjang dari kepatuhan pasien dalam mengikuti arahan dokter dan pada program rehabilitasi medik khususnya dalam beberapa minggu pascaoperasi (Wu et al., 2019). Prosedur fisioterapi pascaoperasi dapat meningkatkan kekuatan dan kecepatan berjalan, mencegah komplikasi, meningkatkan mobilitas pasien, dan memberikan edukasi terkait latihan lanjutan setelah pasien keluar dari rumah sakit (Coulter et al., 2013). Setelah menjalani operasi THA, pasien akan menjalani serangkaian protokol rehabilitasi medik yang telah ditetapkan oleh rumah sakit berupa sejumlah kunjungan untuk memantau kondisi pascaoperasi dan melakukan rehabilitasi fungsi gerak pasien. Rumah sakit memberikan rekomendasi dan panduan untuk tiap fase rehabilitasi, yang dapat sedikit berbeda pada tiap rumah sakit. Sebagai contoh Tabel 2.1 menjelaskan protokol pascaoperasi di Wexner Medical Center, The Ohio State University, Amerika Serikat (Davis, 2018).

SEKOLAH PASCASARJANA

Tabel 2.1. Contoh Protokol Pascaoperasi THA.

	Fase	
	I: Hari ke-1 pasca-operasi s.d. lepas alat bantu (0-6 minggu)	II: lepas alat bantu s.d. bebas nyeri pada ADL (6-12 minggu)
Tujuan	1. melindungi jaringan untuk penyembuhan	1. Restorasi Passive dan Active ROM
	2. Kontrol rasa nyeri dan edema	2. Peningkatan kekuatan proximal hip musculature
	3. Pencegahan Deep Vein Thrombosis	3. Normalisasi pelvis
	4. Meningkatkan ROM yang bebas nyeri	4. Normalisasi gaya berjalan di area komunitas
	5. Normalisasi otot	5. Toleransi ADL tanpa nyeri atau batasan
	6. Dapat bergerak tanpa alat bantu jalan	
	7. ADL yang independen	
Tindakan Pencegahan	1. Pendekatan anterior: a. No hip extension > 20° b. No hip external rotation > 50°	Ekstensi dari pencegahan pada fase I
	2. Pendekatan posterior: a. No hip flexion > 90° b. No hip internal rotation atau adduction lebih dari posisi netral	
	3. Secara umum: a. Menahan beban sesuai toleransi, dengan penggunaan alat bantu (tongkat, walker) b. Tidak boleh menyilangkan kaki c. Gerakan membungkuk/mengangkat yang baik (punggung lurus dan lutut ditekuk) d. Jaga panggul tetap di atas lutut saat duduk (hindari kursi yang terlalu dalam)	
Progresi alat bantu	1. Walker -> tidak terlalu membatasi (23issal: tongkat), atau tanpa alat bantu	
	2. 2 -> 1 -> 0 kruk yang dapat ditoleransi	
Kriteria Ambulasi Komunitas tanpa alat bantu	1. ROM panggul yang memadai untuk pola berjalan normal/bebas nyeri (10° hip extension)	
	2. 60 detik <i>single leg stance</i> (SLS) tanpa kompensasi (panggul turun, badan condong) atau rasa sakit	
	3. Pola berjalan normal tanpa alat bantu	

Tabel 2.1. Contoh Protokol Pascaoperasi THA (lanjutan).

	Fase	
	I: Hari ke-1 pasca-operasi s.d. lepas alat bantu (0-6 minggu)	II: lepas alat bantu s.d. bebas nyeri pada ADL (6-12 minggu)
ROM/ peregangan	1. Passive ROM (tanpa nyeri): hip flexion, ekstensi ke netral jika ada kontraktur	1. Mobilisasi jaringan lunak dan sendi
	2. Passive ROM lembut, flexion Active-Assistive ROM terlentang sesuai pedoman	2. Active ROM posisi naik ke sepeda
	3. ROM dalam posisi seperti naik ke sepeda	3. Rekomendasi ke terapi pijat jika tumbuh jaringan lunak yang disfungsi atau iritasi
	4. Mobilisasi jaringan lunak dan bekas luka setelah jahitan menutup	4. Secara kontinyu menilai aktivitas pasien di luar sesi terapi fisik
Kontrol neuromuscular	Bagian ini prioritas no 1 -> jangan lanjut ke penguatan sebelum aktivasi otot dan kontrol isolasi dinormalisasi	
	Glute sets, quad set, transverse abdominis, hamstrings, dilakukan terlentang atau dengan kaki ditekuk	
Latihan terapis	<u>Latihan Awal:</u> Isometrics, Short/Long Arc Quad, berjinjit, Straight Leg Raise, weight shifting	<u>Latihan Awal:</u> mini squats hingga flexion 70°, melangkah ke samping, SLS pada permukaan tidak stabil
	<u>Latihan Lanjut:</u> syarat: pola gerak normal, rasa nyeri reaktif minimal Straight Leg Raise (flexi, abduksi, extensi), naik/turun tangga	<u>Latihan Lanjut:</u> leg press, mini squats, terapi air dapat dilakukan setelah luka bedah sembuh dan atas rekomendasi dokter
Latihan kardiovaskular		Latihan bersepeda selama 30 menit, dilanjutkan dengan penambahan beban
Kriteria untuk Lanjut ke Fase berikutnya	1. Pola gerak normal di area rumah tanpa alat bantu 2. Rasa nyeri dan bengkak reaktif minimal atau bebas nyeri saat ADL dan latihan fisik 3. Aktivasi dan isolasi otot normal 4. SLS > 20 detik tanpa panggul turun	1. Hip ROM yang simetris dan bebas nyeri 2. Kekuatan lower extremity yang baik (skor 4 dari 5) 3. DL Simetris hingga 70° tekukan lutut 4. Gerakan yang bagus sesuai tes Forward Step Down 5. Pola jalan normal dalam area komunitas

2.2.2 Pengukuran Hasil dan Skor Fungsi Panggul

Pengukuran dan penilaian pada hasil operasi THA dapat menggunakan beberapa macam skor (Weick et al., 2020) seperti:

1. *Harris Hip Score (HHS)*

HHS mengukur nyeri, fungsionalitas, ketiadaan deformitas, dan rentang gerak (ROM). Domain nyeri mengukur tingkat keparahan nyeri dan pengaruhnya terhadap aktivitas dan kebutuhan akan pengobatan nyeri (44 poin). Domain fungsi menilai aktivitas sehari-hari (menggunakan tangga, transportasi umum, duduk, memakai sepatu dan kaus kaki) serta gaya berjalan (47 poin). Deformitas mengevaluasi fleksi panggul, adduksi, rotasi internal, dan perbedaan panjang ekstremitas (4 poin). Rentang gerak mengukur fleksi panggul, abduksi, rotasi eksternal dan internal, dan adduksi (5 poin). Versi modifikasi dari HHS (*Modified Harris Hip Score-MHHS*) sekarang tersedia yang hanya mengukur komponen nyeri dan fungsionalitas saja. HHS dan MHHS memiliki rentang nilai 0 sampai 100 (pada MHHS skor akhir dikalikan dengan 1,1 agar tetap menghasilkan total 100 poin). Semakin tinggi skor, semakin baik capaian hasil operasi dengan klasifikasi: < 70 bernilai buruk (*Poor*), 70-80 bernilai cukup (*Fair*), 80-90 bernilai baik (*Good*), dan 90-100 bernilai sempurna (*Excellent*).

2. *Hip disability and Osteoarthritis Outcome Score (HOOS)*

HOOS mengukur opini pasien dewasa tentang panggul dan disabilitasnya dari 5 domain: (1) nyeri, (2) gejala lainnya, (3) fungsionalitas pada *Activities of Daily Living* (ADL), (4) fungsionalitas pada olahraga dan rekreasi, dan (5) kualitas hidup terkait panggul. Terdapat total 40 pertanyaan dengan rentang nilai 0 sampai 100. Setiap pertanyaan dibuat dalam bentuk 5 skala Likert dengan skor 0 sampai 4 (tidak ada gejala, ringan, sedang, berat, ekstrim). Beberapa variasi dari HOOS adalah:

- a. HOOS-JR (*HOOS-Joint Replacement Short Form*): diisi oleh pasien yang menjalani THA untuk mengukur domain nyeri dan aktivitas harian.
- b. HOOS-PS (*HOOS-Physical Function Short Form*): mengukur fungsi fisik untuk menggali opini pasien terkait kesulitan yang dialami karena masalah

panggul.

- c. HOOS-12 (*HOOS-12 items*): mengukur domain nyeri, fungsionalitas, dan kualitas hidup.

HOOS memiliki batas minimal perbedaan skor yang dianggap bermakna (MCID) yakni antara 9,6 pada domain ADL dan 16,2 pada domain QoL.

3. *Oxford Hip Score (OHS)*

OHS mengukur tingkat nyeri (6 pertanyaan) dan fungsionalitas (6 pertanyaan) pada panggul terkait aktivitas sehari-hari seperti berjalan, mengenakan pakaian, dan lain-lain. Rentang skor OHS antara 0 sampai 48 (terburuk sampai terbaik) untuk setiap pertanyaan dengan total skor antara 0 sampai 48 (paling sulit sampai tidak ada kesulitan). Beberapa penelitian mengusulkan klasifikasi skor > 41 sebagai sangat baik, 34-41 baik, 27-33 cukup, dan <27 buruk. Sejumlah penelitian terbaru menunjukkan bahwa skor OHS pascaoperasi yang lebih besar dari 37,5 dikategorikan sebagai operasi yang sukses. Sebuah studi di tahun 2015 menyatakan MCID sebesar 5 poin setelah THA.

2.2.3. *Explainable Artificial Intelligence (XAI)*

1. Alasan Dibutuhkan XAI

Keberadaan kecerdasan buatan dan *machine learning* (AI/ML) memberikan peluang untuk meningkatkan kualitas analisis data dengan kemampuannya untuk menemukan pola dan struktur pada data yang memiliki ukuran besar dan dimensi yang tinggi dengan menghasilkan sebuah model. Meskipun demikian, sifat model yang dihasilkan oleh *machine learning* secara mendasar bersifat *black box*, di mana output keputusan (berupa klasifikasi atau regresi) tidak dapat dipahami bagaimana keputusan tersebut dapat dihasilkan dan dipercaya atau divalidasi oleh manusia. Solusi berbasis AI/ML yang dibutuhkan pun sekarang telah meningkat ke solusi (model) yang *explainable* dan *interpretable* atau bersifat transparan, khususnya dalam bidang kesehatan di mana diagnosis dan prognosis yang dihasilkan dari AI/ML harus dapat memenuhi persyaratan medis, etik, serta legalitas dalam menghasilkan keputusan yang masuk akal layaknya justifikasi ahli seorang dokter (Amann et

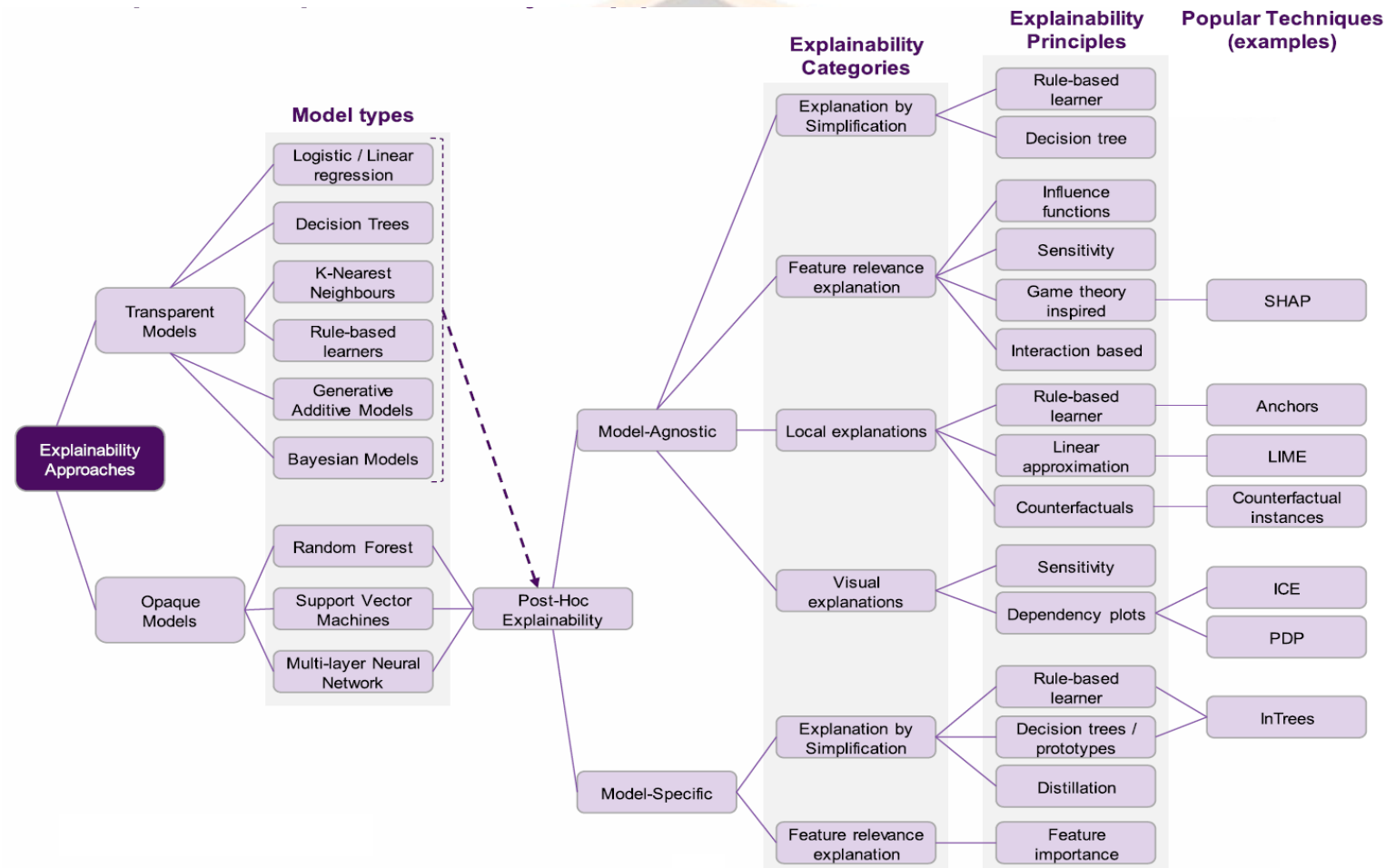
al., 2020; Belle & Papantonis, 2021; Chaddad et al., 2023). Model dengan karakteristik yang dapat dijelaskan tersebut disebut *Explainable Artificial Intelligence* (XAI). Berikut alasan yang mendorong kebutuhan akan interpretabilitas dan penjelasan dari model AI/ML (Doshi-Velez & Kim, 2017; Molnar, 2025):

- a. Keingintahuan manusia dan proses pembelajaran dalam hal pencarian makna dalam kehidupan
 - b. Pencapaian tujuan sains yakni ekstraksi pengetahuan
 - c. Sebagai alat bantu untuk identifikasi langkah-langkah keamanan pada sistem AI/ML yang dibuat untuk menyelesaikan permasalahan riil
 - d. Deteksi bias pada model
 - e. Peningkatan penerimaan sosial
2. Taksonomi Teknik XAI

Belle dan Papantonis (Belle & Papantonis, 2021) telah memetakan taksonomi XAI yang dapat memberikan gambaran ikhtisar terhadap teknik XAI seperti Gambar 2.2. Secara umum ada 2 (dua) pendekatan dari teknik XAI, yakni langsung menggunakan model yang transparan dan secara intrinsik *interpretable* atau menggunakan model tidak transparan yang dilengkapi dengan penjelasan setelah model dihasilkan (disebut *post-hoc*).

Transparansi dari sebuah model menunjukkan tingkat pemahaman manusia terhadap cara kerja internal dari model yang mencakup 3 (tiga) dimensi (Lipton, 2017):

- a. **Simulatabilitas:** transparansi tingkat pertama dan mengacu pada kemampuan model untuk disimulasikan oleh manusia
- b. **Dekomposabilitas:** transparansi tingkat kedua yang menyatakan kemampuan model untuk didekomposisi menjadi bagian-bagian (input, parameter, dan komputasi) yang diberi penjelasan
- c. **Transparansi Algoritma:** transparansi tingkat ketiga dan menyatakan kemampuan algoritma untuk dapat menjelaskan prosedur yang dilalui untuk memproses input menjadi output



Gambar 2.2. Taksonomi XAI: Peta Pendekatan *Explainability* (Belle dan Papantonis, 2021)

Tabel 2.2 membandingkan berbagai tipe model dari Gambar 2.2 berdasarkan transparansinya, dengan Regresi Linier/Logistik, Pohon Keputusan, *K-nearest neighbors*, *Rule-based learners*, dan model *Bayesian* dikatakan sebagai model yang transparan dan memiliki karakteristik yang mudah dipahami sehingga tidak memerlukan penjelasan tambahan setelah model dihasilkan. Di sisi lain, model yang bersifat *black box*, seperti *Random Forest*, *Support Vector Machine*, dan *multi-layer neural network* membutuhkan penjelasan tambahan setelah model dihasilkan atau disebut *post-hoc*.

3. Ruang Lingkup Penjelasan

Ruang lingkup penjelasan pada XAI dapat bersifat global atau lokal. Penjelasan yang bersifat global mengevaluasi seluruh basisdata yang tersedia, baik yang melibatkan keseluruhan maupun sebagian kecil fitur, untuk memberikan penjelasan terkait cara kerja model. Penjelasan model yang bersifat global secara praktek sulit dilakukan pada data yang memiliki puluhan, ratusan, bahkan ribuan parameter, misal pada data teks atau citra. Jika menggunakan penjelasan yang sederhana (dengan menggunakan model yang transparan untuk mengimitasi kerja dari model *black box*) maka berpotensi memiliki presisi manusia yang rendah. Model yang sederhana tidak dapat menangkap secara utuh perilaku dari model yang kompleks dan dapat menjerumuskan pengguna pada kesimpulan yang salah (Ribeiro et al., 2018).

Sebaliknya, penjelasan yang bersifat lokal mengevaluasi sebuah prediksi terhadap satu data dan berusaha mengungkap mengapa prediksi tersebut dibuat terhadap data tersebut dengan memberikan relasi sebab-akibat antara input dengan hasil prediksi (Nagahisarchoghaei et al., 2023). Penjelasan dibuat terhadap model yang sangat kompleks dan tetap masih dapat dipahami oleh manusia, namun hanya terbatas pada area spesifik dari ruang input saja.

Tabel 2.2. Perbandingan tipe model berdasarkan transparansi (Belle & Papantonis, 2021)

Model	Simulatabilitas	Dekomposabilitas	Transparansi algoritma	Post-hoc
Regresi Linier/Logistik	Prediktor dapat dibaca manusia dan interaksi antar-prediktor dibuat minimal	Terlalu banyak interaksi dan prediktor	Variabel dan interaksi terlalu kompleks untuk dianalisis tanpa <i>tool</i> matematika	Tidak perlu
Pohon keputusan	Manusia dapat memahami tanpa latar belakang matematika	Aturan yang dibuat tidak memodifikasi data dan dapat dipahami	Manusia dapat memahami prediksi dengan menelusuri pohon	Tidak perlu
Rule-based learners	Variabel yang dapat dibaca, jumlah aturan masih dapat dikelola	Jumlah aturan terlalu banyak untuk dianalisis	Aturan sangat kompleks sehingga dibutuhkan <i>tool</i> matematika	Tidak perlu
<i>General additive models</i>	Variabel, interaksi, dan fungsi harus dapat dipahami	Interaksi terlalu kompleks untuk disimulasikan	Karena kompleksitasnya, variabel dan interaksinya tidak dapat dianalisis tanpa <i>tool</i> matematika	Tidak perlu
Model Bayesian	Relasi statistik dan variabel harus dapat dipahami oleh pengguna	Relasi melibatkan terlalu banyak variabel	Relasi dan prediktor sangat kompleks sehingga dibutuhkan <i>tool</i> matematika	Tidak perlu
Pohon Ansambel	Tidak tersedia	Tidak tersedia	Tidak tersedia	<i>Feature relevance, model simplification</i>
<i>Support Vector Machine</i>	Tidak tersedia	Tidak tersedia	Tidak tersedia	<i>Feature relevance, model simplification</i>
Jaringan Saraf Tiruan	Tidak tersedia	Tidak tersedia	Tidak tersedia	<i>Feature relevance, visualisasi, model simplification</i>

4. Teknik XAI *Model-agnostic*

Teknik *model-agnostic* memiliki keunggulan dapat diterapkan pada model AI/ML apapun tanpa bergantung pada sifat intrinsik arsitektur model tertentu. Kelebihan dari metode penjelasan yang *model-agnostic* adalah fleksibilitasnya

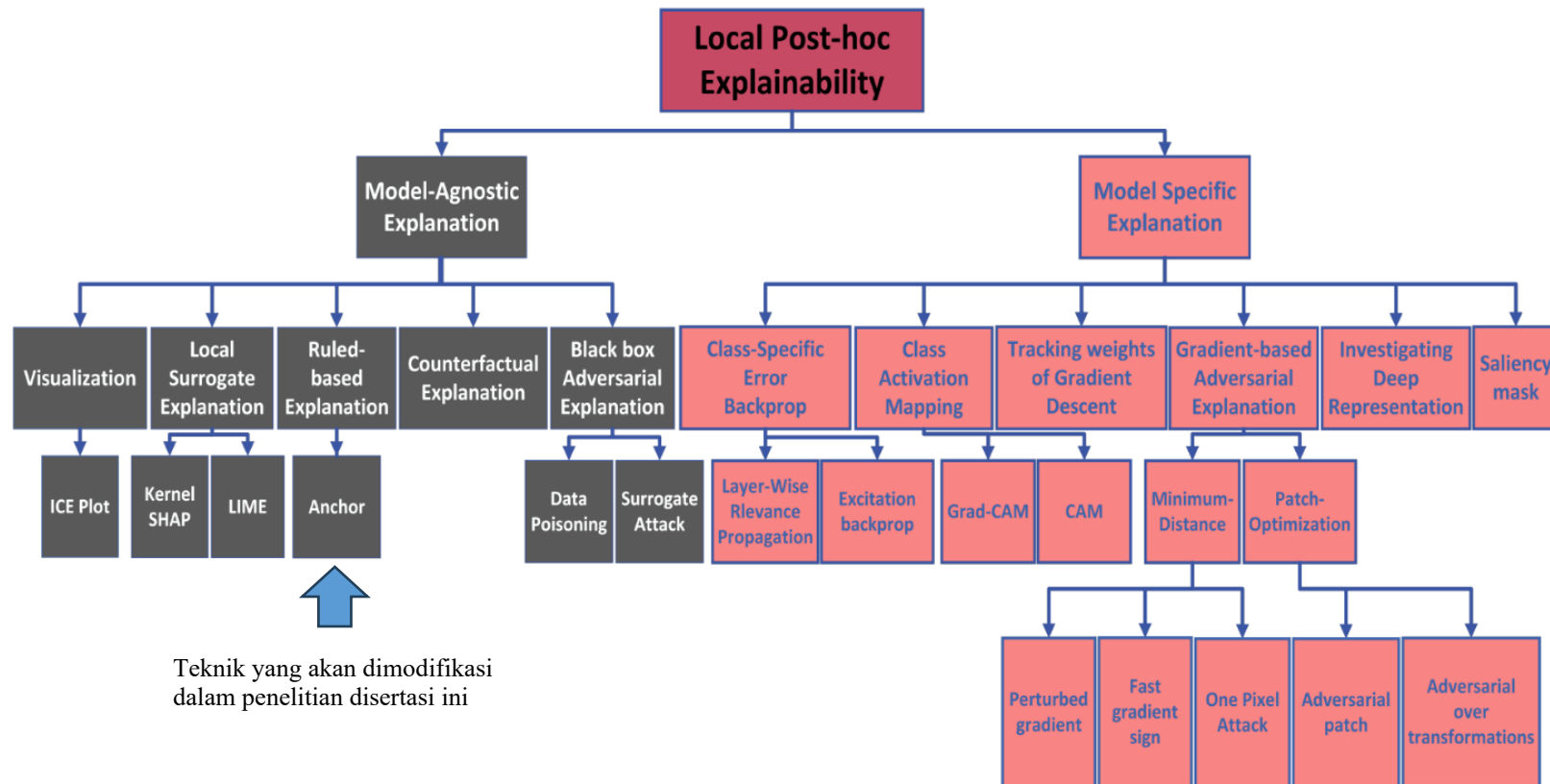
yang memberikan kebebasan bagi developer untuk menggunakan model *machine learning* apapun yang dinilai paling tepat untuk menyelesaikan sebuah permasalahan, tidak harus selalu menggunakan model yang transparan. Fleksibilitas juga berlaku pada metode penjelasan yang ingin digunakan, yang dapat berupa formulasi linier, *feature importance*, atau bentuk penjelasan lain yang dianggap sesuai dengan permasalahan atau target pengguna yang dituju (Ribeiro et al., 2016).

Alternatif dari penjelasan *model-agnostic* adalah menggunakan penjelasan yang *model-specific*, untuk model yang berbasis runut balik dan *deep neural network*. Nagahisarchoghaei melakukan studi lanjutan dan pemetaan yang lebih lengkap terkait evolusi teknik XAI melalui bibliometrik dan analisis jaringan (Nagahisarchoghaei et al., 2023). Gambar 2.3 menunjukkan berbagai metode interpretabilitas yang bersifat lokal yang dapat terdiri atas *model-agnostic* dan *model-specific*. Penelitian disertasi ini akan melakukan pengembangan dan modifikasi pada teknik *Anchors* yang merupakan penjelasan *model-agnostic* yang berbasis aturan.

5. Sifat-sifat dari penjelasan XAI

Penjelasan XAI tidak memiliki kebenaran mutlak, sehingga evaluasi dari penjelasan yang dibangkitkan dilakukan secara kualitatif, dan terkadang secara kuantitatif. Sebuah penjelasan merelasikan nilai fitur dari sebuah *instance* dengan prediksi model dengan cara yang dapat dipahami oleh manusia. Sejumlah sifat-sifat dari sebuah penjelasan yang dibangkitkan adalah sebagai berikut (Molnar, 2025):

- a. **Akurasi**, yakni menilai seberapa baik sebuah penjelasan dapat memprediksi data yang belum pernah dilihat. Akurasi yang tinggi sangat penting jika penjelasan digunakan sebagai pengganti prediksi model *machine learning*. Meskipun demikian, akurasi yang rendah masih dapat diterima jika akurasi model *machine learning* juga rendah dan tujuannya adalah untuk menjelaskan cara kerja dari model *machine learning* tersebut.



Gambar 2.3. Teknik XAI yang bersifat lokal dan *post-hoc* (Nagahisarchoghaci et al., 2023)

- b. **Fidelitas**, yakni menilai seberapa baik penjelasan dapat mengaproksimasi prediksi model *black box*. Fidelitas yang tinggi adalah salah satu komponen terpenting dalam sebuah penjelasan karena menunjukkan kinerja dalam memberikan penjelasan. Beberapa teknik XAI memberikan fidelitas yang bersifat lokal dengan penjelasan yang hanya memiliki aproksimasi baik untuk sekumpulan data atau bahkan hanya satu data individu saja.
- c. **Konsistensi**, yakni menilai seberapa berbeda penjelasan yang dibangkitkan pada berbagai model *machine learning* yang dilatih pada basisdata yang sama. Tidak ada keharusan untuk penjelasan memiliki konsistensi yang tinggi, jika dibangkitkan dari berbagai model *machine learning* yang mungkin saja menggunakan fitur yang berbeda dalam menghasilkan penjelasan.
- d. **Stabilitas**, yakni menilai seberapa mirip penjelasan yang dibangkitkan untuk *instance* yang mirip, dengan menggunakan model *machine learning* yang sama. Berbeda dengan konsistensi yang membandingkan penjelasan pada berbagai model, stabilitas membandingkan penjelasan antar *instance* yang serupa pada sebuah model tertentu.
- e. **Komprehensibilitas** menilai seberapa baik manusia memahami penjelasan yang dibangkitkan. Fitur ini sulit untuk didefinisikan dan diukur, namun beberapa pakar sepakat bahwa hal ini bergantung pada audiens yang menerima penjelasan.
- f. **Kepastian**, yakni mengukur tingkat kepastian dari model *machine learning*. Penjelasan yang dibangkitkan perlu menyatakan tingkat kepercayaan bahwa hasil prediksi yang dibuat adalah benar.
- g. **Derajat Kepentingan**, yakni mengukur seberapa baik penjelasan yang dibangkitkan dapat merefleksikan pentingnya sebuah fitur atau bagian dari penjelasan. Jika sebuah aturan dihasilkan dalam sebuah penjelasan, maka perlu diperjelas kondisi aturan mana yang paling penting.
- h. **Kebaruan**, yakni menilai apakah jika penjelasan diterapkan pada data baru yang berada di luar distribusi data latih maka model *machine learning* bisa jadi kurang akurat dan penjelasan yang dibangkitkan menjadi tidak berguna. Konsep kebaruan berkaitan erat dengan kepastian, yakni semakin tinggi kebaruannya, semakin mungkin model memiliki kepastian yang rendah.

- i. **Keterwakilan** mengukur seberapa banyak *instance* yang dicakup oleh penjelasan yang dibangkitkan. Penjelasan dapat mencakup keseluruhan model (misalnya pada interpretasi bobot pada model regresi linier) atau hanya mewakili satu prediksi individu saja (misalnya pada *Shapley values*).

6. *Anchors*

Anchors adalah teknik penjelasan berdasarkan penyederhanaan yang diusulkan oleh Ribeiro (Ribeiro et al., 2018). Ribeiro juga merupakan peneliti yang mengusulkan teknik penjelasan LIME (*Local Interpretable Model-Agnostic Explanations*) yang merupakan teknik penjelasan paling populer dalam kategori penjelasan berdasarkan penyederhanaan (Belle & Papantonis, 2021). Berbeda dengan LIME yang menggunakan *linear model* yang dibangun di sekitar hasil prediksi model untuk menghasilkan model pengganti, *Anchors* membangun *rule if-then* yang lebih mudah dipahami dalam mengartikan ciri yang terpenting saja untuk menghasilkan penjelasan yang lebih ramping. Secara formal, *Anchors* didefinisikan dengan Persamaan (2.1):

$$\mathbb{E}_{\mathcal{D}_x(z|A)}[1_{\hat{f}(x)=\hat{f}(z)}] \geq \tau, A(x) = 1 \quad (2.1)$$

dengan:

- x adalah *instance* yang akan dijelaskan, misal berupa sebuah baris dalam tabel
- A adalah himpunan predikat, yakni *anchor* atau aturan yang dihasilkan, sedemikian hingga $A(x) = 1$ ketika semua fitur predikat didefinisikan oleh A berkorespondensi dengan nilai fitur x
- f menyatakan model klasifikasi yang akan dijelaskan, misal sebuah jaringan saraf tiruan, yang dapat dipanggil untuk memprediksi label dari *instance* x dan sejumlah tetangganya
- $\mathcal{D}_x(\cdot | A)$ mengindikasikan distribusi dari tetangga x yang sesuai dengan aturan A
- $0 \leq \tau \leq 1$ menyatakan *threshold* presisi dari sebuah *Anchor*. Hanya aturan yang memenuhi fidelitas lokal minimal sebesar τ yang diakui sebagai hasil yang valid

Karena Persamaan (2.1) sulit untuk dihitung secara langsung, maka digunakan definisi probabilitas dengan parameter $0 \leq \delta \leq 1$, di mana sampel data akan terus diambil sampai tercapai tingkat kepercayaan statistik yang cukup dari nilai

presisinya. *Anchor* yang valid adalah yang memenuhi batasan presisi dengan tingkat probabilitas yang tinggi seperti pada Persamaan (2.2):

$$P(\text{prec}(A) \geq \tau) \geq 1 - \delta \text{ dengan } \text{prec}(A) = \mathbb{E}_{\mathcal{D}_x(z|A)}[1_{\hat{f}(x)=\hat{f}(z)}] \quad (2.2)$$

dengan $\mathbb{E}_{\mathcal{D}_x(z|A)}[1_{\hat{f}(x)=\hat{f}(z)}]$ adalah harapan matematis nilai prediksi model $\hat{f}(x)$ sama dengan $\hat{f}(z)$ pada pengambilan sampel z jika diketahui *anchor* A pada distribusi $\mathcal{D}$ dari *instance* x , atau dengan kata lain model selalu memberikan output yang konsisten.

Anchor yang didapat akan diukur tingkat cakupannya (disebut *coverage*) terhadap *instance* di sekitarnya dan memilih *anchor* yang memiliki *coverage* tertinggi. Persamaan (2.3) mendefinisikan *coverage* sebuah *anchor* dan Persamaan (2.4) mendefinisikan *coverage* tertinggi dari sebuah *anchor*.

$$\text{cov}(A) = \mathbb{E}_{\mathcal{D}(x)}[A(z)] \quad (2.3)$$

$$\max_{A \text{ s.t. } P(\text{prec}(A) \geq \tau) \geq 1 - \delta} \text{cov}(A) \quad (2.4)$$

Dengan demikian, proses pencarian difokuskan pada penemuan aturan yang memiliki *coverage* tertinggi dari seluruh aturan yang memenuhi persyaratan, yakni yang memenuhi ambang batas probabilitas *precision*. Sebuah aturan *anchor* dengan banyak predikat akan memiliki *precision* yang lebih tinggi dari aturan dengan sedikit predikat, karena aturan yang “mengikat” lebih banyak predikat pada sebuah *instance* x akan menyusutkan ruang tetangga yang dievaluasi, menghasilkan *instance* yang identik dengan x , memberikan prediksi yang sama, dan memiliki *precision* aturan bernilai 1. Di sisi lain, aturan yang memiliki banyak predikat menandakan *anchor* yang terlalu spesifik dan hanya berlaku pada sedikit *instance*, sehingga terjadi *trade-off* antara *precision* dan *coverage*. Pendekatan *Anchors* memiliki 4 (empat) komponen utama, yaitu:

- a. **Pembangkitan Kandidat:** pada ronde pertama dihasilkan satu kandidat untuk setiap fitur x . Pada ronde berikutnya, kandidat terbaik dari ronde sebelumnya diperluas dengan menambahkan satu fitur predikat lain yang belum terdapat dalam himpunan kandidat.
- b. **Identifikasi Kandidat Terbaik:** Menggunakan *Multi-Armed Bandit* KL-LUCB (Kaufmann dan Kalyanakrishnan, 2013) untuk memilih kandidat aturan mana yang paling baik dapat menjelaskan x .

- c. **Validasi Presisi Kandidat:** Perulangan untuk terus melakukan pengambilan sampel data jika aturan kandidat masih belum memenuhi ambang batas τ .
- d. **Algoritma Modifikasi Pencarian Beam:** Ketiga komponen di atas diramu dalam sebuah modifikasi algoritma pencarian Beam yang merupakan algoritma pencarian graf yang memiliki *hyperparameter* B , yakni *Beam Width*, yang merepresentasikan B kandidat anchor terbaik pada setiap ronde.

Kode 2.1 merupakan *pseudocode Anchor* untuk mengidentifikasi kandidat terbaik dengan pendekatan *Greedy Search*, sedangkan Kode 2.2 adalah *pseudocode* dari algoritma modifikasi *Beam Search* yang digunakan *Anchors* untuk memilih dari beberapa kandidat *rule anchors* yang memiliki tingkat cakupan paling tinggi terhadap hasil prediksi (Ribeiro et al., 2018). Kode 2.1 menerjemahkan Persamaan (2.2) di mana ketika sebuah aturan dipilih oleh KL-LUCB sebagai sebuah *anchor*, maka aturan tersebut harus memenuhi kriteria *precision* yang ditetapkan. Jika sebuah aturan A terpilih namun $prec_{lb} < \tau$ dan $prec_{ub} > \tau$, maka pengambilan sampel dari $\mathcal{D}(\cdot | A)$ terus dilakukan sampai ada kepercayaan bahwa A adalah *anchor*, yakni $prec_{lb} > \tau$, atau bukan, yakni $prec_{ub} < \tau$.

Algorithm 1 Identifying the Best Candidate for Greedy	
function GenerateCands($\mathcal{A}, c$) $\mathcal{A}_r = \emptyset$ for all $A \in \mathcal{A}; a_i \in x, a_i \notin A$ do if $cov(A \wedge a_i) > c$ then $\mathcal{A}_r \leftarrow \mathcal{A}_r \cup (A \wedge a_i)$ return $\mathcal{A}_r$ function BestCand($\mathcal{A}, \mathcal{D}, \epsilon, \delta$) initialize $prec, prec_{ub}, prec_{lb}$ estimates $\forall A \in \mathcal{A}$ $A \leftarrow \argmax_A prec(A)$ $A' \leftarrow \argmax_{A' \neq A} prec_{ub}(A', \delta)$ while $prec_{ub}(A') - prec_{lb}(A) > \epsilon$ do sample $z \sim \mathcal{D}(z A), z' \sim \mathcal{D}(z' A')$ update $prec, prec_{ub}, prec_{lb}$ for A and A' $A \leftarrow \argmax_A prec(A)$ $A' \leftarrow \argmax_{A' \neq A} prec_{ub}(A')$ return A	{only high coverage} {Add as potential anchor} {Candidate anchors for next round} { δ implicit below } {Sample more}

Kode 2.1. Fungsi GenerateCands dan BestCand untuk identifikasi kandidat terbaik *anchors*

Pada Kode 2.1 fungsi *GenerateCands*, c adalah parameter nilai *coverage* yang akan dibandingkan dengan nilai *coverage* baru dengan penambahan atribut a_i pada kandidat *anchor* A . Fungsi *BestCand* mencari kandidat terbaik dari kembalian fungsi *GenerateCands* yang memenuhi batas *precision* aturan dengan rata-rata tertinggi, A , dan batas atas tertinggi, A' , dan memperbarui nilai batas atas $prec_{ub}$ dan

bawah $prec_{lb}$ dengan mengambil sampel z dan z' dari $\mathcal{D}(z|A)$ dan $\mathcal{D}(z'|A')$ sampai batas bawah dari A lebih tinggi dari batas atas A' dengan toleransi $\epsilon \in [0,1]$.

Algorithm 2 Outline of the Beam Search <pre> function BeamSearch($f, x, \mathcal{D}, \tau$) hyperparameters B, ϵ, δ $A^* \leftarrow \text{null}, \mathcal{A}_0 \leftarrow \emptyset$ {Set of candidate rules} loop $\mathcal{A}_t \leftarrow \text{GenerateCands}(\mathcal{A}_{t-1}, \text{cov}(A^*))$ $\mathcal{A}_t \leftarrow B - \text{BestCand}(\mathcal{A}_t, \mathcal{D}, B, \delta, \epsilon)$ {LUCB} if $\mathcal{A}_t = \emptyset$ then break loop for all $A \in \mathcal{A}_t$ s.t. $prec_{lb}(A, \delta) > \tau$ do if $\text{cov}(A) > \text{cov}(A^*)$ then $A^* \leftarrow A$ return A^* </pre>
--

Kode 2.2. Fungsi Beam Search untuk pencarian aturan *anchors*

Pada Kode 2.2, parameter f adalah *black-box classifier* yang digunakan untuk pemanggilan model prediksi terhadap *instance* tetangga dari x , yang merupakan sebuah *instance* pada distribusi $\mathcal{D}$, dan τ adalah level presisi yang diinginkan. Notasi A^* menyatakan aturan yang memiliki *highest true precision* dari seluruh aturan yang memenuhi level presisi. Algoritma memilih B kandidat terbaik dengan toleransi $\epsilon \in [0,1]$ dan mengembalikan himpunan $\mathcal{A}$ berukuran B yang merupakan aproksimasi- ϵ dari A^* dengan probabilitas tinggi. Pemilihan kandidat terbaik diimplementasikan dengan algoritma penyelesaian permasalahan *Multi Arm Bandit* (MAB) versi KL-LUCB (Kaufmann dan Kalyanakrishnan, 2013).

Untuk menunjukkan bagaimana *anchor* dapat memberikan penjelasan yang dapat dipahami oleh pengguna, Gambar 2.4 mengilustrasikan sebuah *instance* dari basisdata pasien yang mendapat prediksi “SKOR_FUNGSIONAL = 60” untuk nilai skor fungsional panggul setelah operasi THA berdasarkan atribut seperti Usia, Jenis Kelamin, Kategori Pembiayaan, Faktor Komorbid, dan lain-lain.

Atribut	Nilai
Usia	75
Jenis Kelamin	Perempuan
Kategori Pembiayaan	BPJS
Komorbid Diabetes	Tidak
Komorbid Malnutrisi	Ya
Komorbid Dementia	Tidak
Atribut Lainnya	
SKOR_FUNGSIONAL	60

Gambar 2.4. Ilustrasi hasil prediksi dari sebuah *instance* pada basisdata pasien

Penjelasan yang diberikan oleh *anchor* terhadap hasil prediksi tersebut adalah:

IF Komorbid Malnutrisi = Ya **AND** Kategori Pembiayaan = BPJS **THEN**
 PREDICT SKOR_FUNGSIONAL = 60 **WITH** PRECISION 97% **AND** COVERAGE 15%

Penjelasan yang diberikan menunjukkan atribut mana yang berkontribusi terhadap hasil prediksi, dalam hal ini adalah jenis komorbid malnutrisi dan kategori pembiayaan asuransi BPJS. *Anchor* juga memberikan informasi bahwa ketepatan presisi dari penjelasan adalah sebesar 97% dan aturan penjelasan tersebut berlaku pada 15% ruang distribusi (*perturbation space* atau $\mathcal{D}$).

Keuntungan dari *anchors* yang utama adalah memiliki output yang jelas dan mudah dipahami, bahkan bagi pengguna awam, karena berbentuk aturan *if-then* yang memiliki informasi derajat kepentingan dengan adanya *coverage* dari aturan tersebut. Keuntungan yang kedua adalah *anchors* bekerja dengan baik pada model prediksi yang non-linier dalam suatu domain distribusi sebuah *instance*. Cara kerja *anchors* yang berbasis *reinforcement learning* (dengan penerapan *Multi-Armed Bandit* atau MAB) lebih baik ketimbang penjelasan yang berbasis model pengganti dan kecil kemungkinannya untuk *underfit*. Selain itu, penggunaan MAB juga menjadikan algoritma *anchors* sangat efisien karena dapat diparalelkan di mana MAB mendukung *batch sampling* (Molnar, 2025).

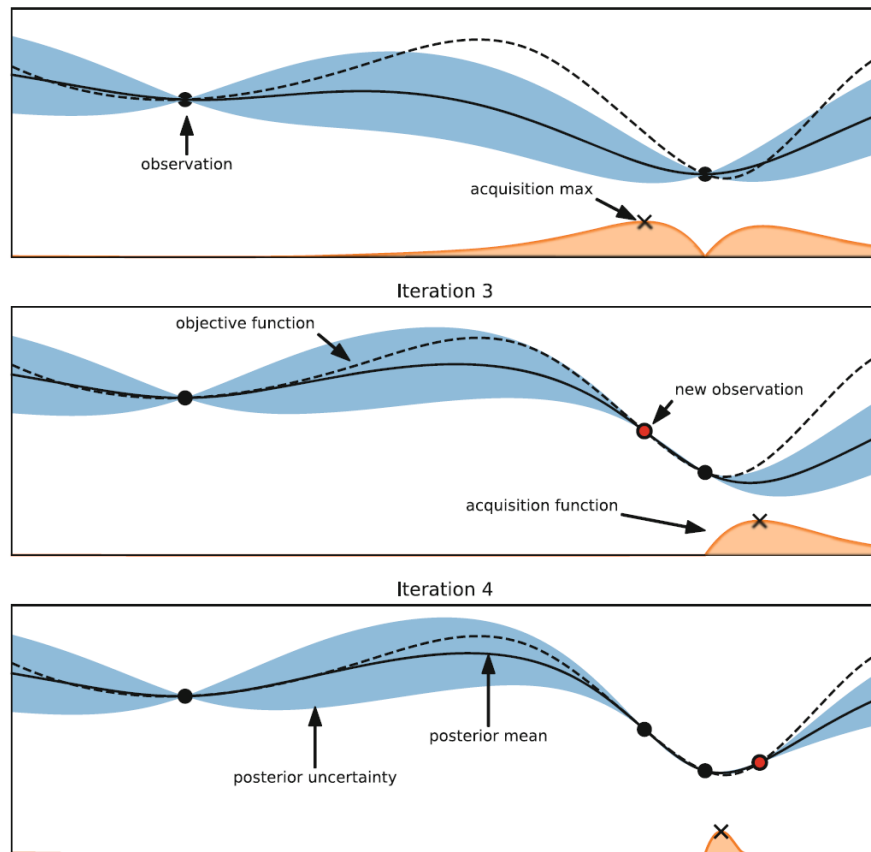
Keterbatasan dari *anchors* adalah tingginya tingkat konfigurasi yang perlu dilakukan agar dapat memberikan hasil yang bermakna, seperti pengaturan *hyperparameter* lebar *beam* pada *Beam Search* dan nilai *threshold* presisi, serta desain fungsi *perturbation* untuk mencari daerah tetangga di sekitar *instance* yang akan diberi penjelasan. Kekurangan lainnya adalah pada banyak skenario dibutuhkan diskretisasi data agar hasil penjelasan tidak terlalu spesifik, memiliki cakupan rendah, dan tidak berkontribusi dalam memberikan penjelasan model. Sebuah penelitian (Delaunay dkk., 2020) mengusulkan metode pemilihan fungsi diskretisasi yang lebih tepat untuk digunakan pada *Anchors*. Selain itu, membangun *anchors* membutuhkan banyak pemanggilan ke fungsi model *machine learning* yang berpotensi mempengaruhi kinerjanya (Molnar, 2025).

2.2.4. Bayesian Optimization

Bayesian Optimization merupakan salah satu *framework* optimisasi global dari fungsi objektif *black box* yang kompleksitasnya tinggi dan tidak memiliki bentuk tertentu (Garnett, 2023; Mockus, 1989). Bayesian Optimization terdiri atas 2 (dua) elemen kunci: fungsi pengganti probabilistik dan fungsi akuisisi untuk menentukan titik mana yang akan dievaluasi. Gambar 2.5 mengilustrasikan bagaimana Bayesian Optimization mengoptimasi sebuah fungsi 1-dimensi.

Bayesian Optimization adalah salah satu teknik Hyperparameter Optimization (HO) yang membangun sebuah model probabilitas dari fungsi objektif dan menggunakannya untuk memilih *hyperparameter* yang paling menjanjikan untuk dievaluasi dalam fungsi objektif yang sesungguhnya (Koehrsen, 2018). Dengan menggunakan pendekatan Bayesian, Bayesian Optimization mencatat hasil evaluasi sebelumnya yang digunakan untuk membangun model probabilitas yang memetakan *hyperparameter* ke skor probabilitas dari fungsi objektif: $P(\text{score}|\text{hyperparameters})$ dan disebut fungsi pengganti. Fungsi ini jauh lebih mudah dioptimasi daripada fungsi objektif dan metode Bayesian bekerja dengan menemukan kumpulan *hyperparameter* berikutnya untuk mengevaluasi fungsi objektif sesungguhnya dengan memilih *hyperparameter* yang berkinerja terbaik pada fungsi pengganti.

Beberapa contoh fungsi akuisisi yang dapat digunakan antara lain adalah *probability of improvement*, *expected improvement* (Jones et al., 1998), *Bayesian expected losses*, *upper confidence bounds*, dan *Thompson sampling* (Russo et al., 2018; Thompson, 1933). Salah satu implementasi yang menggunakan teknik Bayesian Optimization adalah *Sequential Model-Based Optimization* (SMBO) (Kode 2.3) (Bergstra et al., 2011). Beberapa contoh fungsi akuisisi yang dapat digunakan antara lain adalah *probability of improvement*, *expected improvement* (Jones et al., 1998), *Bayesian expected losses*, *upper confidence bounds*, dan *Thompson sampling* (Russo et al., 2018; Thompson, 1933). Salah satu implementasi yang menggunakan teknik Bayesian Optimization adalah *Sequential Model-Based Optimization* (SMBO) (Kode 2.3) (Bergstra et al., 2011).



Gambar 2.5. Ilustrasi Bayesian Optimization pada fungsi 1-dimensi (garis putus-putus) menggunakan *surrogate function* Gaussian (garis hitam), dengan area biru adalah daerah *uncertainty* dan memaksimalkan fungsi akuisisi (kurva oranye)

SMBO (f, M_0, T, S)	
1	$\mathcal{H} \leftarrow \emptyset,$
2	For $t \leftarrow 1$ to T ,
3	$x^* \leftarrow \operatorname{argmin}_x S(x, M_{t-1}),$
4	Evaluate $f(x^*),$ {Expensive step}
5	$\mathcal{H} \leftarrow \mathcal{H} \cup (x^*, f(x^*)),$
6	Fit a new model M_t to $\mathcal{H},$
7	return $\mathcal{H}$

Kode 2.3. Algoritma SMBO untuk *Bayesian Optimization*

dengan $\mathcal{H}$ adalah riwayat observasi pasangan (*hyperparameter*, skor), T adalah iterasi maksimum, f adalah fungsi objektif sejati (*True Objective Function*), M adalah fungsi pengganti (*Surrogate Function*) yang diperbarui jika ada sampel baru yang ditambahkan, S adalah fungsi akuisisi (*Acquisition Function*), dan x^* adalah *hyperparameter* terpilih berikutnya yang akan dievaluasi.

Salah satu implementasi *Bayesian Optimization* adalah pustaka *bayes-opt* yang berbasiskan Python yang menggunakan fungsi pengganti berupa proses Gaussian

(Nogueira, 2014). Pada setiap langkah optimasi, sebuah proses Gaussian, sebagai fungsi pengganti, disesuaikan dengan sampel yang sudah diketahui (titik yang sudah pernah dieksplorasi), dan distribusi posterior yang dikombinasikan dengan strategi seperti *Upper Confidence Bound*, *Probability of Improvement*, atau *Expected Improvement*, sebagai fungsi akuisisi, akan digunakan untuk menentukan titik berikutnya yang akan ditelusuri. *Bayesian Optimization* dengan proses Gaussian dinilai dapat memberikan konvergensi yang cepat dan tepat digunakan pada *hyperparameter* bertipe kontinyu dan berjumlah kurang dari 10 (Bischl et al., 2023; L. Yang & Shami, 2020). Untuk mengoptimisasi sebuah fungsi dengan *bayes-opt* perlu dibuat sebuah variabel yang menginstansiasi class *BayesianOptimization* dengan memberikan parameter seperti fungsi yang akan dioptimasi, serta batas atas dan batas bawah dari nilai variabel yang dioptimasi. Kode 2.4 menunjukkan contoh potongan kode untuk optimasi sebuah fungsi *black box f*, dengan penetapan batas atas dan bawah *pbounds*, jumlah *init_points=2* dan jumlah iterasi optimasi=3.

```
from bayes_opt import BayesianOptimization
# Bounded region of parameter space
pbounds = {'x': (2, 4), 'y': (-3, 3)}

optimizer = BayesianOptimization(
    f=black_box_function,
    pbounds=pbounds,
    random_state=1,
)
optimizer.maximize(init_points=2, n_iter=3)
```

Kode 2.4. Contoh potongan kode penggunaan *Bayesian Optimization*

2.2.5. Metrik Evaluasi Kinerja

Evaluasi kinerja algoritma *machine learning* dilakukan dengan menggunakan metrik seperti Akurasi, Spesifisitas, Sensitivitas (atau *Recall*), Presisi, *F-Score*, *Area Under the Receiver Operating Curve* (AUROC), dan *Brier Score*. Metrik Akurasi (Persamaan 2.5) sering digunakan untuk mengukur kinerja sebuah model namun dapat memberikan informasi yang tidak valid pada dataset yang tidak seimbang, sehingga diperlukan metrik kinerja lain seperti Spesifisitas (Persamaan (2.6), yakni kemampuan model untuk dapat mengidentifikasi seluruh sampel yang negatif sejati), Sensitivitas (atau *Recall* (Persamaan (2.7)), yang merupakan kemampuan model untuk dapat mengidentifikasi seluruh sampel yang positif

sejati), dan Presisi (atau *Positive Predictive Value* (Persamaan (2.8)), yakni ketepatan model dalam mengidentifikasi sampel yang positif). Metrik F-Score (Persamaan (2.9)) merupakan rata-rata harmonis yang menunjukkan *trade-off* antara Sensitivitas dan Presisi, sedangkan AUROC yang dihitung menggunakan Sensitivitas dan *False Positive Rate* (Persamaan (2.10)) mengukur seberapa efisien model yang dihasilkan yang ditunjukkan dengan kemampuan model untuk mendiskriminasi antara kelas positif dan negatif. *Brier Score* (Persamaan (2.11)) mengukur ketepatan probabilitas hasil prediksi untuk menilai kualitas kalibrasi model prediksi yang dihasilkan, khususnya pada permasalahan multi kelas dengan nilai semakin mendekati 0 menunjukkan semakin baik kalibrasi antara hasil prediksi dengan nilai sesungguhnya.

$$Akurasi = \frac{TP + TN}{(TP + TN + FP + FN)} \quad (2.5)$$

$$Spesifisitas = TNR = \frac{TN}{(TN + FP)} \quad (2.6)$$

$$Sensitivitas = Recall = TPR = \frac{TP}{(TP + FN)} \quad (2.7)$$

$$Presisi = PPV = \frac{TP}{(TP + FP)} \quad (2.8)$$

$$F - Score = \frac{2 * Presisi * Recall}{Presisi + Recall} \quad (2.9)$$

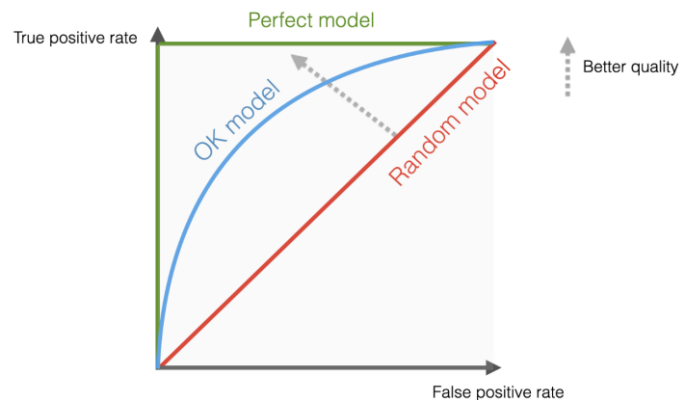
$$FPR = \frac{FP}{(FP + TN)} \quad (2.10)$$

$$Brier\ Score\ multiclass = \frac{1}{N} \sum_{t=1}^N \sum_{i=1}^R (f_{ti} - o_{ti})^2 \quad (2.11)$$

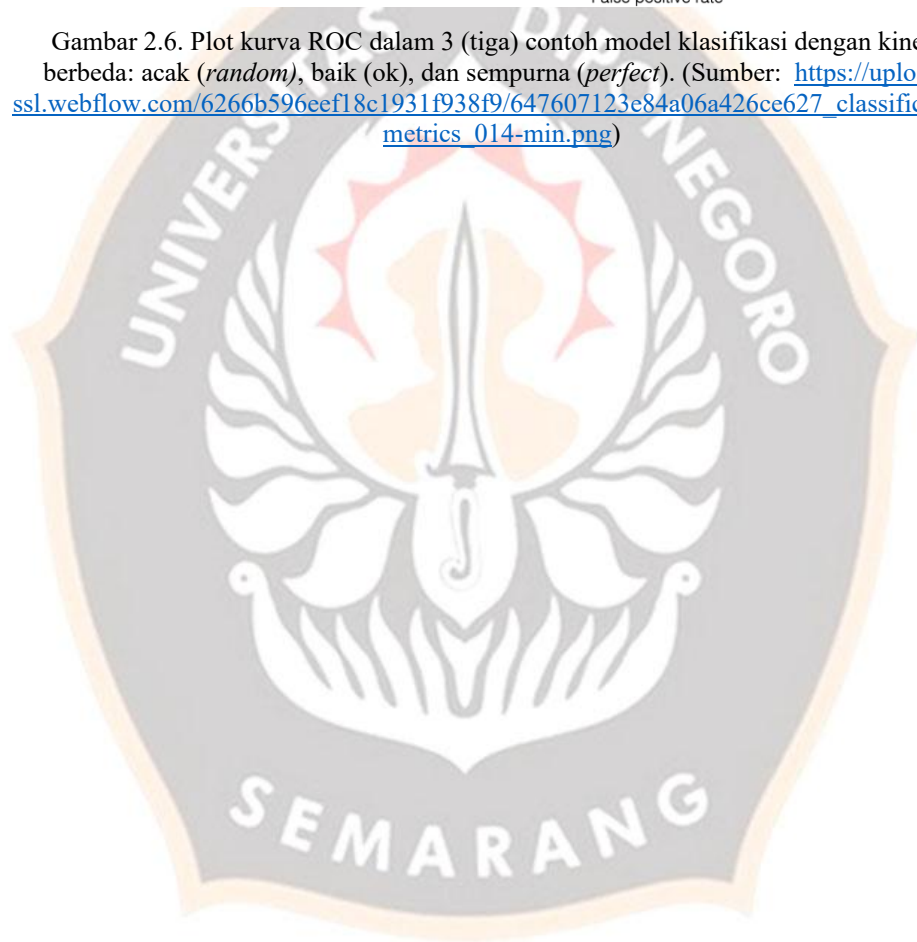
Keterangan:

TP = *True Positive*; TN = *True Negative*; FP = *False Positive*; FN = *False Negative*; TNR = *True Negative Rate*; TPR = *True Positive Rate*; PPV = *Positive Predictive Value*; FPR = *False Positive Rate*; N=jumlah seluruh *instance*; R=jumlah kelas; f_{ti} = probabilitas prediksi untuk kelas i ; o_{ti} =1 jika *instance* t ada di kelas i dan 0 jika bukan.

Gambar 2.6 menunjukkan plot kurva AUROC antara nilai TPR dan FPR. Kurva ROC memplot 2 (dua) parameter, yakni Sensitivitas (*True Positive Rate* atau TPR) dan *False Positive Rate* (FPR) dalam sebuah grafik dan luas area di bawah kurva tersebut disebut sebagai AUROC. Model dengan skor AUROC > 0,8 dikategorikan baik dan skor > 0,9 dikategorikan sangat baik. Model yang sempurna memiliki kurva siku sepanjang sumbu y, sedangkan model acak memiliki kurva 45°. Model yang baik memiliki kurva yang mendekati model sempurna.



Gambar 2.6. Plot kurva ROC dalam 3 (tiga) contoh model klasifikasi dengan kinerja berbeda: acak (*random*), baik (*ok*), dan sempurna (*perfect*). (Sumber: https://uploads-ssl.webflow.com/6266b596eef18c1931f938f9/647607123e84a06a426ce627_classification_metrics_014-min.png)



SEKOLAH PASCASARJANA