

Pengembangan Machine Learning dan Metode Heuristik

Kecerdasan Artifisial(CIF63310 / 2 sks)

Building Up
Noble Future

Outline

- Pendahuluan
- Pendekatan Hybrid dalam AI
- Bagging
- Boosting
- Stacking
- Rule based + ML
- Heuristik + ML
- Contoh-contoh Hybrid

Pendahuluan



- Apa itu Machine Learning dan Heuristik?
 - Machine Learning (ML): Algoritma yang belajar dari data untuk membuat prediksi/keputusan.
 - Heuristik: Metode berbasis aturan atau pengalaman untuk menyelesaikan masalah secara cepat.

Pendekatan Hybrid dalam AI



- Definisi Pendekatan Hybrid
 - Penggabungan berbagai teknik (ML, heuristik, rule-based, dll.) untuk meningkatkan performa atau efisiensi.
 - Mengatasi keterbatasan dari masing-masing metode.

Bagging



- **Teknik Ensemble - Bagging (Bootstrap Aggregating)**

- *Bagging* mengurangi variansi dengan menggabungkan prediksi dari beberapa model independen.
- **Contoh:** *Random Forest* - Bagging dengan *decision trees*.

- **Formulasi:**

$$\hat{f}(x) = \frac{1}{B} \sum_{b=1}^B f_b(x)$$

- $f_b(x)$ adalah model individu, dan B adalah jumlah model.

Boosting



- Boosting adalah teknik machine learning yang menggabungkan model-model lemah untuk membentuk model yang kuat.
- Setiap model lemah (weak learner) diperkuat dengan mempelajari kesalahan dari model sebelumnya.



Tujuan Utama Boosting

1. Memperbaiki kesalahan dari model sebelumnya.
2. Membentuk model yang lebih akurat dengan menggabungkan model lemah.
3. Meningkatkan akurasi tanpa overfitting.

Formulasi Umum Boosting

- Misalkan data pelatihan $D=\{(x_1,y_1),(x_2,y_2),\dots,(x_n,y_n)\}$.
- Model akhir $F(x)$ dibentuk sebagai kombinasi dari beberapa model lemah $h_t(x)$:

$$F(x) = \sum_{t=1}^T \alpha_t \cdot h_t(x)$$

- dimana α_t adalah bobot model pada iterasi t .

Alur dan Formulasi Proses Boosting

1. Inisialisasi bobot awal $w_i = 1/n$.
2. Melatih model lemah $h_t(x)$ pada setiap iterasi t .
3. Hitung error model t :

$$\text{Error}_t = \sum_{i=1}^n w_i \cdot 1(y_i \neq h_t(x_i))$$

di mana $1(y_i \neq h_t(x_i))$ adalah indikator yang bernilai 1, jika $y_i \neq h_t(x_i)$ dan 0 jika sebaliknya.

4. Tentukan bobot model:

$$\alpha_t = \frac{1}{2} \cdot \ln \left(\frac{1 - \text{Error}_t}{\text{Error}_t} \right)$$

5. Update bobot sampel w_i :

$$w_i \leftarrow w_i \cdot \exp(-\alpha_t \cdot y_i \cdot h_t(x_i))$$

6. Gabungkan model:

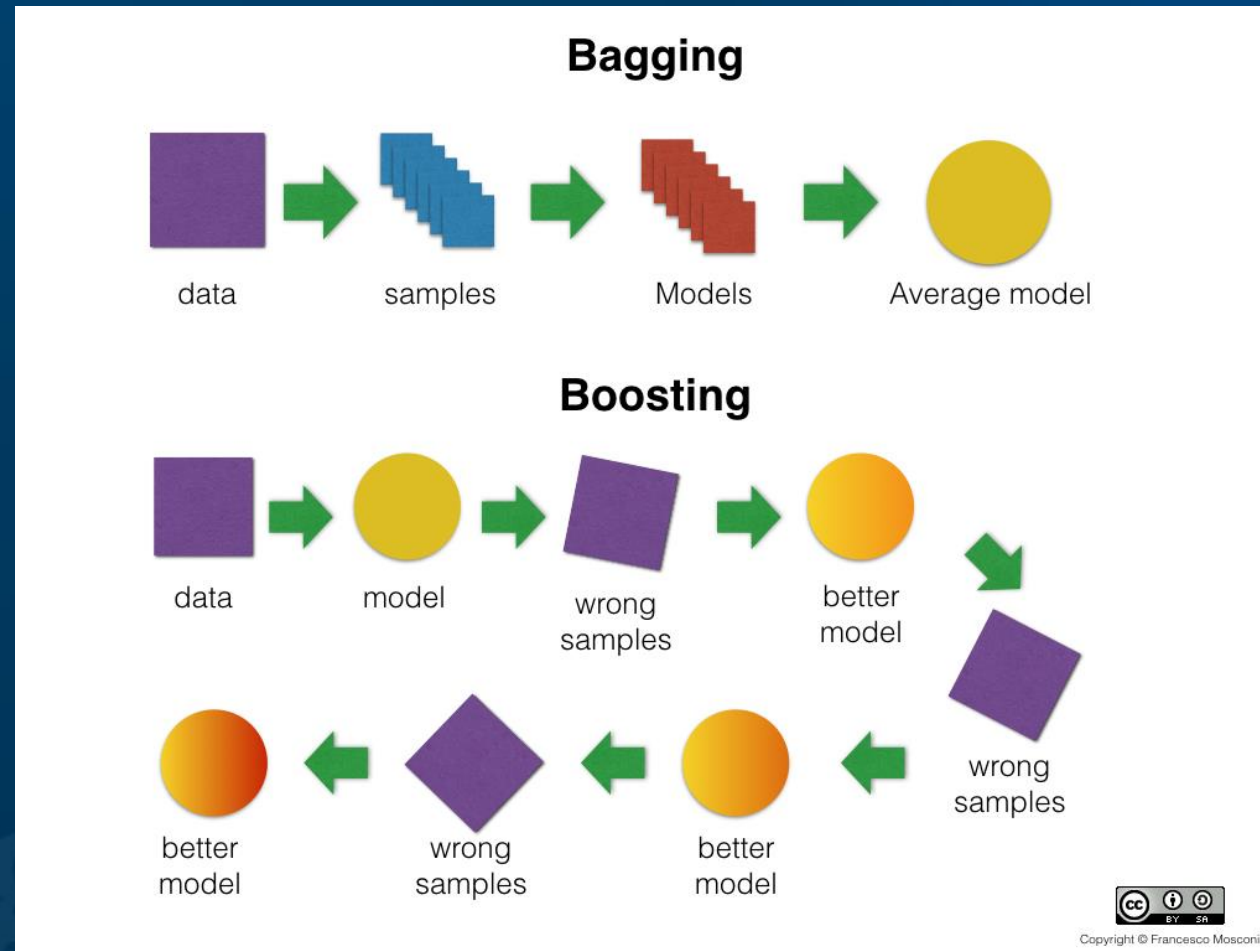
$$F(x) = \sum_{t=1}^T \alpha_t \cdot h_t(x)$$

Contoh Varian Boosting



1. AdaBoost: Menggunakan bobot untuk sampel salah prediksi.
2. Gradient Boosting: Meminimalkan kesalahan residu.
3. XGBoost: Implementasi Gradient Boosting yang dioptimalkan.

Bagging vs Boosting



Stacking



- **Teknik Ensemble - Stacking**

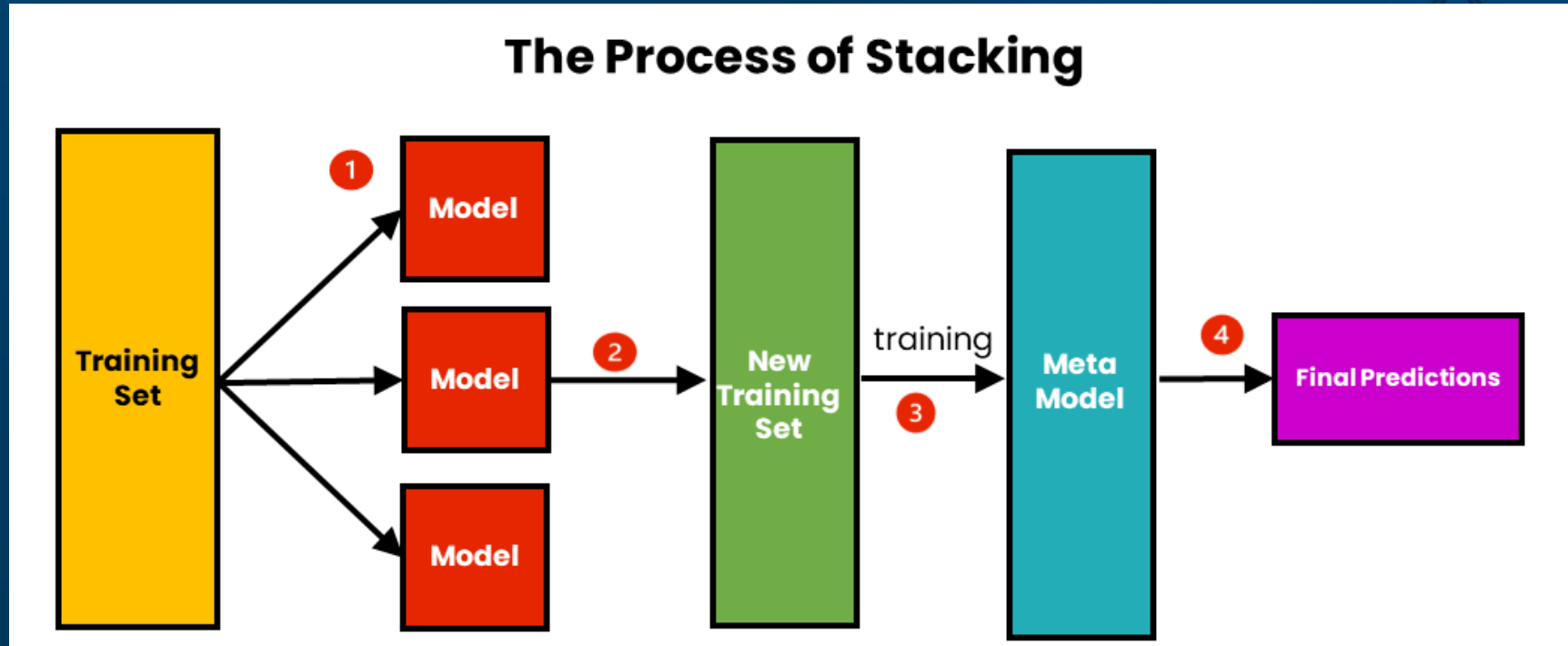
- *Stacking* melibatkan pelatihan beberapa model (level 1), dan hasilnya menjadi input ke model meta (level 2).
- Model meta biasanya melakukan kombinasi linear, seperti regresi linear atau logistik.

- **Formulasi:**

$$\hat{y} = g(f_1(x), f_2(x), \dots, f_n(x))$$

- Di mana g adalah model meta, dan $f_i(x)$ adalah model level ke- i .

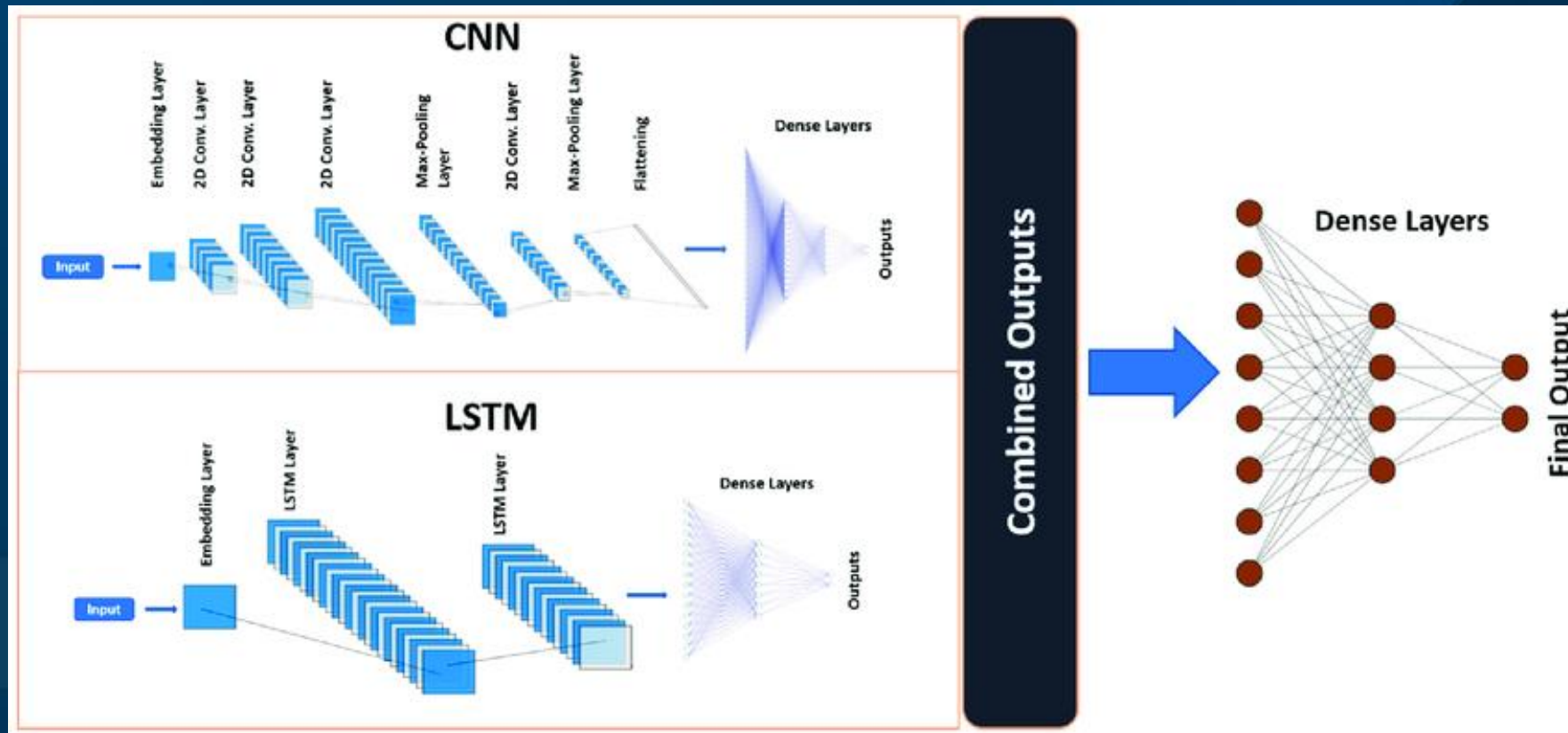
Stacking



Contoh Stacking: CNN + LSTM



- Kombinasi *Convolutional Neural Network* (CNN) untuk ekstraksi fitur dan *Long Short-Term Memory* (LSTM) untuk analisis urutan.
- Digunakan pada data time-series atau video.



Gabungan Rule-Based dan Machine Learning



- Menggabungkan model berbasis aturan dengan ML untuk meningkatkan performa dan interpretabilitas.
- **Contoh:** Pendeteksian penipuan yang menggabungkan aturan berbasis domain dan pembelajaran mesin.

Contoh Rule based + ML (Pendeteksian penipuan)

Misalnya, kita memiliki data transaksi yang berisi beberapa fitur:

- Jumlah Transaksi
- Waktu Transaksi
- Negara Asal Transaksi
- Riwayat Kredit

Langkah 1: Aturan Berbasis Domain



Sebagai langkah awal, kita menetapkan beberapa aturan berbasis domain untuk mendeteksi kemungkinan adanya penipuan. Misalnya:

- 1. Jumlah Transaksi Tinggi:** Jika jumlah transaksi melebihi \$10,000, tambahkan skor risiko sebesar 0.3.
- 2. Transaksi Tengah Malam:** Jika waktu transaksi antara jam 12:00 AM - 4:00 AM, tambahkan skor risiko sebesar 0.2.
- 3. Transaksi dari Negara Berisiko Tinggi:** Jika transaksi dilakukan dari negara yang dianggap berisiko tinggi, tambahkan skor risiko sebesar 0.4.

Misalkan kita memiliki transaksi dengan rincian sebagai berikut:

- Jumlah Transaksi = \$12,000
- Waktu Transaksi = 2:30 AM
- Negara Asal Transaksi = Negara Berisiko Tinggi

Skor Risiko Berdasarkan Aturan Berbasis Domain:

Aturan 1: 0.3 (karena transaksi lebih dari \$10,000)

Aturan 2: 0.2 (karena transaksi terjadi pada tengah malam)

Aturan 3: 0.4 (karena berasal dari negara berisiko tinggi)

Total skor risiko berbasis aturan = $0.3+0.2+0.4=0.9$.

Langkah 2: Model Machine Learning



- Model machine learning dilatih pada data historis untuk memprediksi kemungkinan adanya penipuan berdasarkan beberapa fitur seperti riwayat kredit, frekuensi transaksi, dll. Misalkan model ML ini memberikan probabilitas 0.75 untuk transaksi tersebut (yaitu, kemungkinan bahwa transaksi ini adalah penipuan).

Langkah 3: Kombinasi Aturan Berbasis Domain dan Model Machine Learning



Untuk menggabungkan hasil dari model berbasis aturan dan model ML, kita bisa menggunakan pendekatan rata-rata berbobot, di mana kita memberikan bobot tertentu pada kedua skor. Misalkan kita memberi bobot 0.6 untuk skor berbasis aturan (karena aturan berbasis domain dianggap cukup andal) dan 0.4 untuk skor dari model ML.

- **Formula Penggabungan:**

Skor Akhir = $(0.6 \times \text{Skor Aturan}) + (0.4 \times \text{Skor ML})$

Dengan memasukkan skor yang sudah kita hitung:

Skor Akhir = $(0.6 \times 0.9) + (0.4 \times 0.75) = 0.54 + 0.3 = 0.84$

Skor akhir dari 0.84 menunjukkan bahwa transaksi ini memiliki risiko penipuan yang tinggi

Contoh Penggabungan Heuristik dan Machine Learning

- **Heuristik (Greedy Algorithm):**

- Misalnya, ada lima titik pengiriman: A, B, C, D, dan E, dan kita ingin menemukan rute tercepat dari titik A.
- Dari titik A, rute yang paling dekat adalah B (waktu 2 jam).
- Dari B, rute berikutnya adalah C (waktu 3 jam), dan seterusnya.
- Hasil sementara: $A \rightarrow B \rightarrow C \rightarrow D \rightarrow E$.

- **Machine Learning (Random Forest):** Model machine learning digunakan untuk memperkirakan waktu berdasarkan faktor-faktor tambahan.

- **Prediksi:** Misalnya, model machine learning memprediksi waktu perjalanan antara B dan C adalah 4 jam (karena ada lalu lintas padat).
- **Revisi:** Berdasarkan prediksi ini, model memperbaiki rute yang sebelumnya dipilih dengan estimasi waktu yang lebih akurat.

Contoh-contoh Pendekatan Hybrid dalam AI



- Ensemble Learning: Gabungan model untuk akurasi lebih tinggi (bagging, boosting).
- Neuro-Fuzzy Systems: Kombinasi jaringan saraf dan logika fuzzy.
- Genetic Algorithms + Neural Networks: Algoritma genetika untuk optimasi jaringan saraf.
- Hybrid Deep Learning: Kombinasi CNN dan LSTM untuk analisis data sekuensial.

Keunggulan dan Tantangan Pendekatan Hybrid



- Keunggulan:
 - Fleksibilitas menangani berbagai data dan masalah.
 - Mampu menangkap kompleksitas masalah.
- Tantangan:
 - Komputasi lebih tinggi dan integrasi kompleks.

Aplikasi Dunia Nyata dalam Pendekatan Hybrid

- Medis: Diagnosis kanker (deep learning + decision tree).
- Finansial: Deteksi penipuan dengan ensemble learning.
- Transportasi: Perencanaan rute menggunakan algoritma genetika dan ML.
- Robotika: Navigasi robot dengan reinforcement learning dan heuristik.

Praktik: Membangun Model Hybrid



- Langkah-langkah Implementasi:
 1. Tentukan masalah dan data yang tersedia.
 2. Pilih metode gabungan sesuai kebutuhan.
 3. Tuning parameter dan evaluasi performa.
 4. Validasi hasil pada data uji.



Kiat dan Tantangan dalam Pengembangan Hybrid AI

- Tips:
 - Pahami kekuatan dan kelemahan masing-masing metode.
 - Gunakan ensemble untuk menangani ketidakseimbangan data.
 - Rancang pipeline pengujian yang baik.
- Tantangan:
 - Interpretabilitas model yang kompleks dan biaya komputasi lebih tinggi.

Kesimpulan

- Pendekatan hybrid dalam AI menawarkan solusi yang kuat untuk masalah kompleks.
- Dapat diterapkan dalam berbagai bidang, tetapi memerlukan perencanaan dan tuning yang matang.



Diskusi dan Tanya Jawab



- Ajakan untuk berdiskusi mengenai tantangan dalam implementasi hybrid AI di dunia nyata.
- Ajukan project yang memungkinkan untuk penggunaan model hybrid AI

SEKIAN



**MERDEKA
BELAJAR**

Building Up
Noble Future