

Pengembangan dan Aplikasi IoT

Data Analitik pada IoT



Fakultas Informatika
Telkom University

Data Analitik

Model Pembelajaran Mesin

- Random Forest

- Bayesian Networks

- Convolutional Neural Networks

- Recurrent Neural Network

Penerapan Model Pembelajaran

References

Data Analitik ?



(a) Representasi data analitik



(b) Representasi data analitik

Figure 1: Data analitik

Mengapa IoT memerlukan data analitik

1. Data analitik merupakan metode untuk menganalisis data mentah
 - membuat kesimpulan dari informasi
2. Data analitik bertujuan untuk mencari kejadian
 - biasanya dari sekumpulan data
3. Data IoT bersumber dari beragam alat dan tempat
4. Memerlukan data fusion

Fungsi dari data analitik [1]

- ▶ Preprocessing
 - Filter event yang tidak penting
 - Denaturasi
 - Ekstraksi fitur
 - Transformasi data ke bentuk yang lebih cocok
 - Menambahkan atribut data
- ▶ Alerting
 - Memeriksa data
- ▶ Windowing
 - Menggeser rangkaian peristiwa
 - Mencari data yang diinginkan

► Joins

- Menggabungkan multiple data
- Menghasilkan data baru

► Errors

- Menghubungkan data-data untuk mencari error
- Mencari error pada stream data

► Databases

- Berinteraksi dengan data warehouse

► Temporal events and patterns

- Rangkaian kejadian yang diamati
- Memiliki tujuan untuk mengawasi event of interest

► Tracking

- Kapan dan dimana sesuatu berada
- Kapan dan dimana sesuatu terjadi
- Kapan sesuatu tidak ada dan dimana harus mendapatkannya

► Trands

- Mendeteksi kejadian dalam serial waktu
- Menemukan sebuah pola berdasarkan histori

► Deep analytics pathway

- Menentukan aksi lanjut yang memerlukan proses detail
- Mendeteksi adanya suatu kejadian yang sangat langka pada lingkungan

► Signaling

- Mengembalikan aksi menuju perangkat sensor
- Melakukan komunikasi dengan perangkat sensor

► Control

- Mekanisme untuk mengontrol perangkat analitik
- Manajemen sistem

Terdapat dua tipe sistem pembelajaran [1]

- ▶ Supervised

Data pembelajaran untuk berlatih disediakan diberikan label terkait dengan setiap entri.

misal data pembelajaran adalah sebuah gambar buah yang sudah dilabeli nama buah.

- ▶ Unsupervised

Data pembelajaran untuk berlatih tidak disediakan label.

menggunakan model matematika. Tipikal ini biasanya digunakan untuk menemukan kelompok benda serupa.

Tiga fundamental kegunaan Pembelajaran mesin [1]

1. Klasifikasi

- Memprediksi nilai diskrit
- Tipe supervised
- Memiliki dua varian
 - ▶ Binomial ← dua kategori
 - ▶ Multi kelas ← lebih dari dua kategori
- Memiliki dua klasifer
 - ▶ Linear classifier ← fungsi yang digunakan untuk pemisah antar kelas adalah garis lurus
 - ▶ Nonlinear classifier ← fungsi yang digunakan untuk pemisah antar kelas adalah polinomial

2. Regresi

- Memprediksi nilai kontinu
- Beberapa teknik regresi
 - ▶ Least squares method
 - ▶ Linear regression
 - ▶ Logistic regression

3. Deteksi Anomali

- Melihat barisan trend data untuk memprediksi kejadian

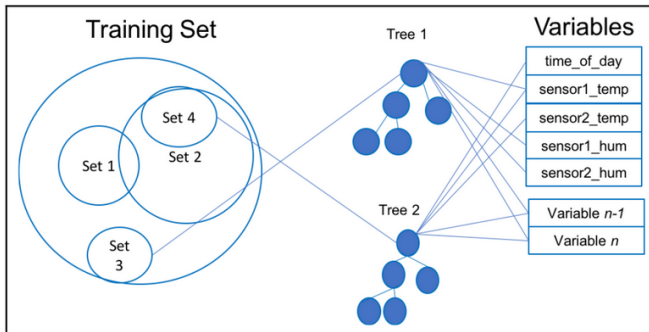


Figure 2: Model random forest [1]

Algoritma

1. Decision tree learning
2. Bagging
3. Random Forest
4. Extra Trees

Kelebihan random forest

- ▶ mampu mengklasifikasi data yang memiliki atribut yang tidak lengkap,
- ▶ dapat digunakan untuk klasifikasi dan regresi akan tetapi tidak terlalu bagus untuk regresi,
- ▶ lebih cocok untuk pengklasifikasian data serta dapat digunakan untuk menangani data sampel yang banyak

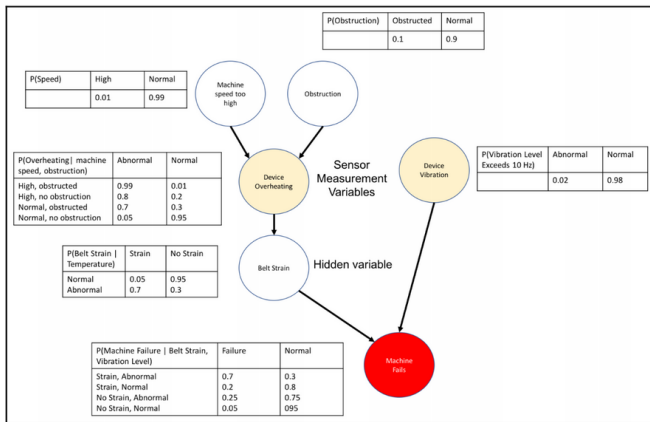


Figure 3: Model bayesian network [1]

Bayesian Network

- ▶ Ideal untuk memprediksi kemungkinan dari beberapa faktor penyebabnya
 - Dengan adanya gejala, dapat digunakan untuk menghitung kemungkinan adanya berbagai penyakit

Kelebihan bayesian network

- ▶ Interpolation
- ▶ Language
mempunyai bahasa tersendiri untuk menetapkan hal-hal yang prior dan posterior. Hal ini secara signifikan membantu pada saat menyelesaikan bagian yang sulit dari sebuah solusi
- ▶ Intuitions
melibatkan prior dan integration, dua aktivitas yang berguna secara luas.

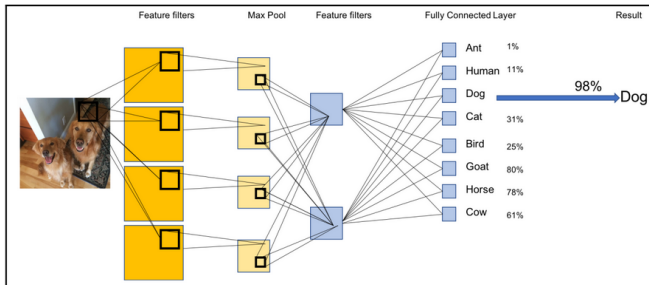


Figure 4: Contoh CNN [1]

Convolutional Neural Networks

- ▶ Ideal untuk klasifikasi gambar
- ▶ Pada IoT diterapkan untuk rekognisi visual
- ▶ Mendukung pada IoT security

Teknik

1. First layer and filters
mengidentifikasi elemen gambar
2. Max pooling and subsampling
mengembalikan nilai maksimal dari masing-masing grid gambar
3. Hidden layers and formal description on forwarding propagation

Kelebihan CNN

- ▶ Meminimalkan komputasi dibandingkan dengan jaringan neural biasa
- ▶ Konvolusi sangat menyederhanakan komputasi tanpa kehilangan esensi data
- ▶ Hebat dalam menangani klasifikasi gambar
- ▶ Menggunakan pengetahuan yang sama di semua lokasi gambar

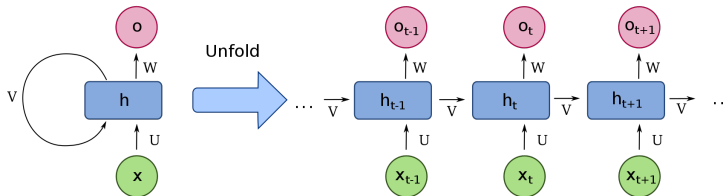


Figure 5: Model RNN

Recurrent Neural Networks

- ▶ Keluaran dari langkah sebelumnya diumpankan sebagai masukan ke langkah saat ini
- ▶ RNN memiliki "memori" yang mengingat semua informasi tentang apa yang telah dihitung

Kelebihan RNN

- ▶ **Memori Jangka Pendek Panjang**
Sebuah RNN mengingat setiap informasi sepanjang waktu. Ini berguna dalam prediksi deret waktu hanya karena fitur untuk mengingat masukan sebelumnya juga. Ini disebut
- ▶ Jaringan neural berulang bahkan digunakan dengan lapisan konvolusional untuk memperluas lingkungan piksel efektif.

Kasus terbaik untuk Random Forest [1]

- ▶ Mendeteksi anomali
- ▶ Sistem dengan ribuan kategori dan ratusan input
- ▶ Regresi dan klasifikasi
- ▶ Menangani tipe data yang tercampur
- ▶ Mengabaikan nilai yang hilang
- ▶ Ukuran linear dengan input

Kurang baik diterapkan pada

- ▶ Ekstraksi fitur
- ▶ Analisis waktu dan urutan

Kasus terbaik untuk Bayesian Networks [1]

- ▶ Data set yang tidak lengkap
- ▶ Pola data streaming
- ▶ Korelasi urutan waktu
- ▶ Data terstruktur
- ▶ Analisis sinyal
- ▶ Model berkembang dengan cepat

Kurang baik diterapkan pada

- ▶ Variabel input independen
- ▶ Berkinerja buruk dengan urutan dimensi data yang tinggi

Kasus terbaik untuk CNN [1]

- ▶ Prediksi suatu objek berdasarkan nilai-nilai sekitarnya
- ▶ Identifikasi pola dan fitur
- ▶ Pengenalan gambar 2D
- ▶ Data Tidak Terstruktur
- ▶ Variabel input mungkin bergantung atau tidak

Kurang baik diterapkan pada RNN

- ▶ Prediksi berbasis waktu dan sekuensial
- ▶ Sistem yang membutuhkan ribuan fitur

Kasus terbaik untuk [1]

- ▶ Prediksi suatu peristiwa berdasarkan urutan
- ▶ Pola data streaming
- ▶ Data seri berkorelasi waktu
- ▶ Data tidak terstruktur
- ▶ Variabel input mungkin bergantung atau tidak

Kurang baik diterapkan pada

- ▶ Analisis gambar dan video
- ▶ Sistem yang membutuhkan ribuan fitur

- [1] P. Lea, *Internet of Things for Architects: Architecting IoT solutions by implementing sensors, communication infrastructure, edge computing, analytics, and security*. Packt Publishing Ltd, 2018.

Thank You!