

# PERTEMUAN 5

## Pemahaman Bisnis Proses Data

# Tujuan Objectif AI dalam Bisnis

**AI** meniru keputusan dan tindakan kecerdasan manusia

Contoh

Autonomous Vehicle

Stock Price Prediction

**Otomatisasi** berfokus pada penyederhanaan tugas instruktif yang berulang

Contoh

Menyalakan lampu menggunakan

smartphone setiap jam 18:00

Alat otomatis pemberi makan hewan peliharaan

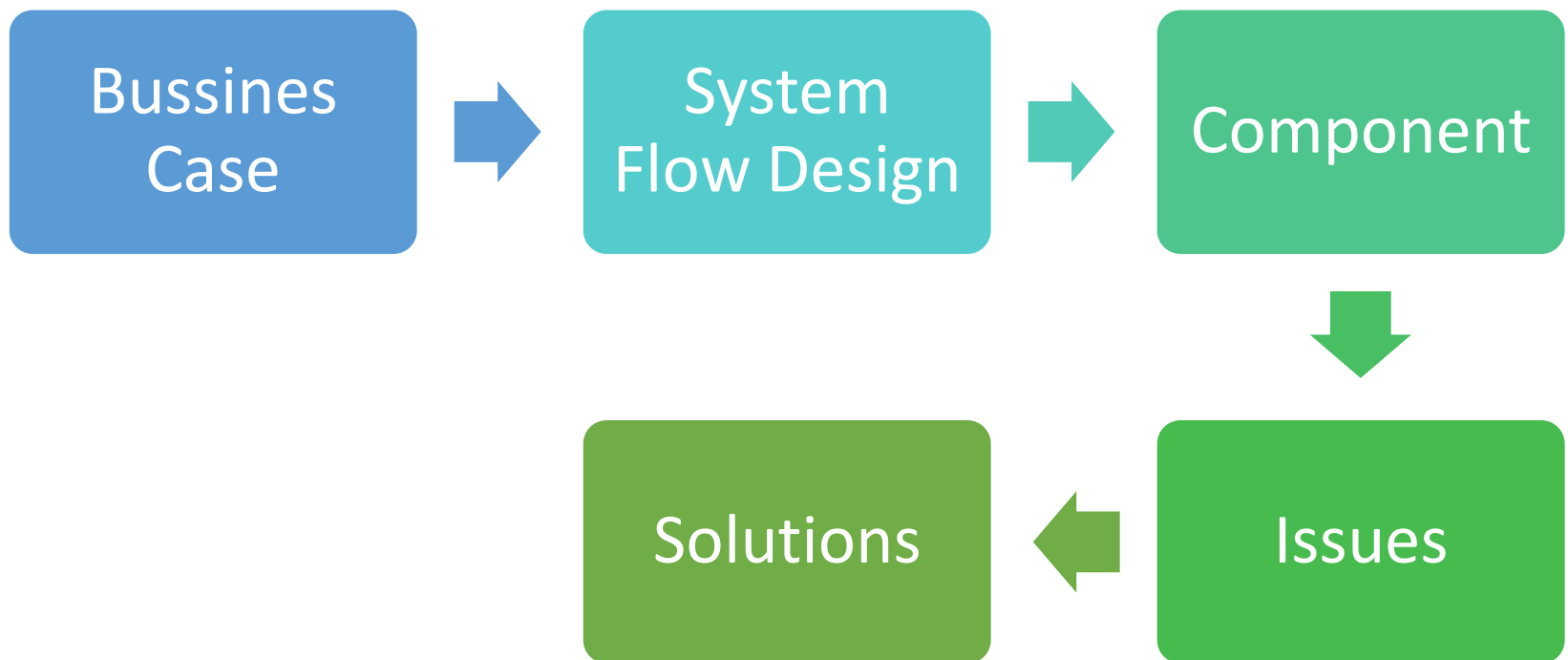
# Identifikasi Problem Machine Learning

Terdapat beberapa sistem pengolahan data, atau sistem pembelajaran yang

memerlukan beberapa tahapan pengolahan yang masing-masing memiliki

keterikatan satu sama lain untuk menghasilkan output.

# Tahapan-tahapan Understanding di dalam AI Business



# Rencana Proyek Data Science

- Setiap proyek dimulai dengan pemahaman bisnis.
- Proyek Data Science merupakan proyek Bisnis, sehingga harus selalu berorientasi pada pencapaian hasil yang berfokus pada bisnis
- Proyek Data Science harus memiliki visi global yang selaras dengan strategi bisnis.
- Sponsor bisnis membutuhkan solusi analitik.

# Tahapan Proyek Data Science



## Analytical Approach

### Contoh bussines understanding:

*Autonomous Vehicle Problem:* Kendaraan autonomous yang belum aman

### Clear question:

Bagaimana cara agar kendaraan autonomous dapat secara aman beroperasi?

### Measurable outcomes:

40% mengurangi kecelakaan lalu lintas yang diakibatkan oleh human error

## *Analytic Approach*

- Setelah masalah bisnis dinyatakan dengan jelas, data scientist dapat menentukan pendekatan analitik untuk memecahkan masalah
- Tahap ini mencakup pengungkapan masalah dalam konteks teknik statistik dan pembelajaran mesin.
- Tahap ini dapat membantu organisasi dalam mengidentifikasi solusi yang paling tepat

# Contoh Analytic approach

"Memprediksi pendapatan perusahaan di kuartal berikutnya?" (Regressi)

"Apakah pasien ini menderita kanker A, kanker B, atau mereka sehat?" (Klasifikasi)

"Apakah ada kelompok pengguna yang tampaknya memiliki kesamaan satu sama lain?" (Clustering)

"Bagaimana saya bisa menargetkan diskon untuk pelanggan tertentu?" (rekomendasi/Personalisasi)



# Analytical Approach

- Pendekatan analitik yang dipilih akan menentukan kebutuhan data.
- Secara khusus, metode analitik yang akan digunakan memerlukan konten, format, dan representasi data tertentu, hal ini juga dipengaruhi oleh problem domain yang dipilih.

# Data Collection

Pada tahap pengumpulan data awal, data scientist dapat melakukan

1. Mengidentifikasi dan mengumpulkan sumber data yang tersedia—

- Terstruktur
- tidak terstruktur
- semi-terstruktur—yang relevan
- dengan domain masalah

# Gathering Data

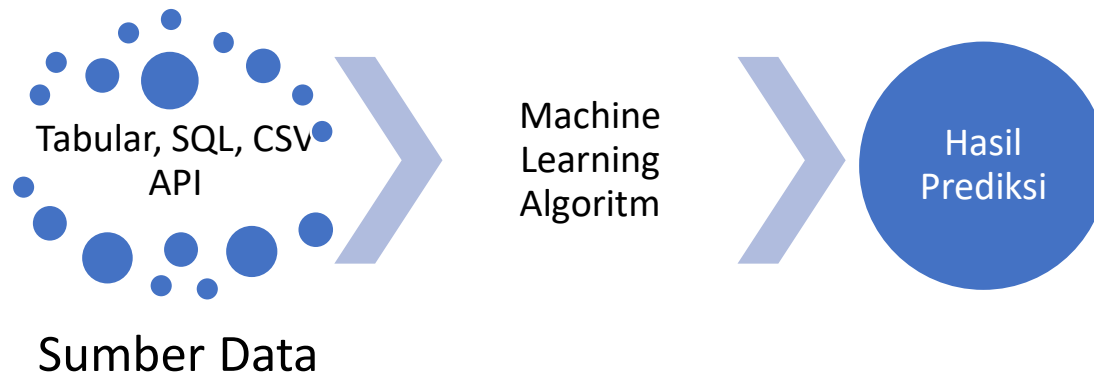
Data dapat dikumpulkan melalui beberapa sumber, seperti

- Data internal perusahaan (excel, database internal, dll)
- Web API, Web scraping
- Kumpulan data melalui data publik
- Kumpulan data melalui data terbuka

# Studi Kasus (Prediksi Harga Saham)

- Pasar saham dikenal sebagai tempat di mana orang dapat berinvestasi dan mendapatkan margin dari investasi tersebut dengan memprediksi harga saham.
- Meskipun sangat sulit untuk memprediksi harga secara tepat.
- Bagaimana kita bisa membuat mesin memiliki kecerdasan untuk dapat memprediksi nilai saham seperti halnya broker di pasar saham?

# System Flow (Prediksi Harga Saham)



- Pada system ini dibutuhkan source data dalam berbagai tipe sumber, tabular, sql, csv, atau API
- Machine learning dibutuhkan untuk mempelajari nilai numerik data-data tersebut
- Hasil prediksi harga saham pada waktu yang ditentukan dapat ditampilkan oleh system berdasarkan data-data lampau

# Permasalahan

- Bagaimana cara agar data dapat di consume oleh algoritma machine learning ? *Solution : Data Cleansing, Data Pre-Processing, Data Understanding*
- Algoritma / solusi machine learning yang perlu diterapkan agar kita bisa mendapatkan prediksi harga saham yang akurat *Solution : Pemilihan algoritma machine learning dengan akurasi terbaik dengan metrik yang sesuai*

# Komponen-Komponen Pendukung

**Data Understanding      Data Cleansing,      Data Pre-processing,      Data**

**Pilihan Machine Learning Algorithm sebagai predictor (Regression)**

- *Auto Regressive Moving Average*
- *Linear Regression*
- *K Nearest Neighbor*
- *Recurrent Neural Network (Long Short Term Memory)*

# Solution : Data cleansing, Data Pre-processing, Data Understanding

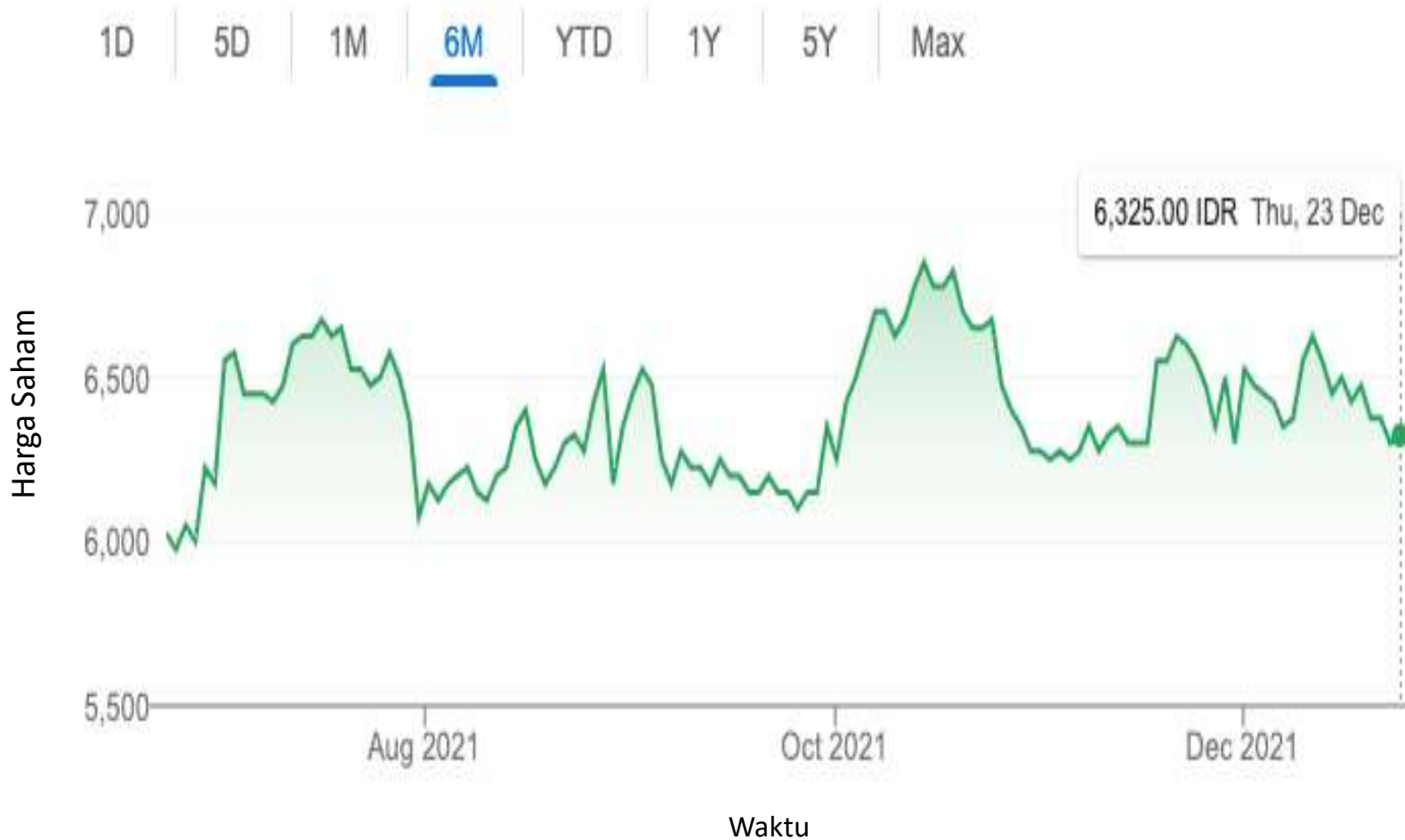
Biasanya suatu data saham memiliki kolom sebagai berikut

	Date	Open	High	Low	Last	Close	Total Trade	Quantity	Turnover (Lacs)
0	2018-10-08	208.00	222.25	206.85	216.00	215.15		4642146.0	10062.83
1	2018-10-05	217.00	218.60	205.90	210.25	209.20		3519515.0	7407.06

- Date, High, low, open, close, Trade Quantity, Turnover
- Date' sebagai fitur indeks. '
- High menunjukkan nilai tertinggi hari itu.
- Low menunjukkan yang terendah.
- Open adalah Harga pembukaan
- Close adalah penutupan untuk Tanggal tersebut.
- Trade Quantity adalah jumlah Saham perusahaan yang diperdagangkan pada tanggal tersebut.
- Turnover adalah omset perusahaan tertentu pada tanggal tersebut



# Data Understanding (Visualisasi)



# Algoritma Regresi

- Auto Regressive Moving Average
- Linear, Lasso, Ridge Regression
- Decision Tree Regression, Adaboost Regression,
- Recurrent Neural Network

# Pengukuran Error

*Regression* Metrics :

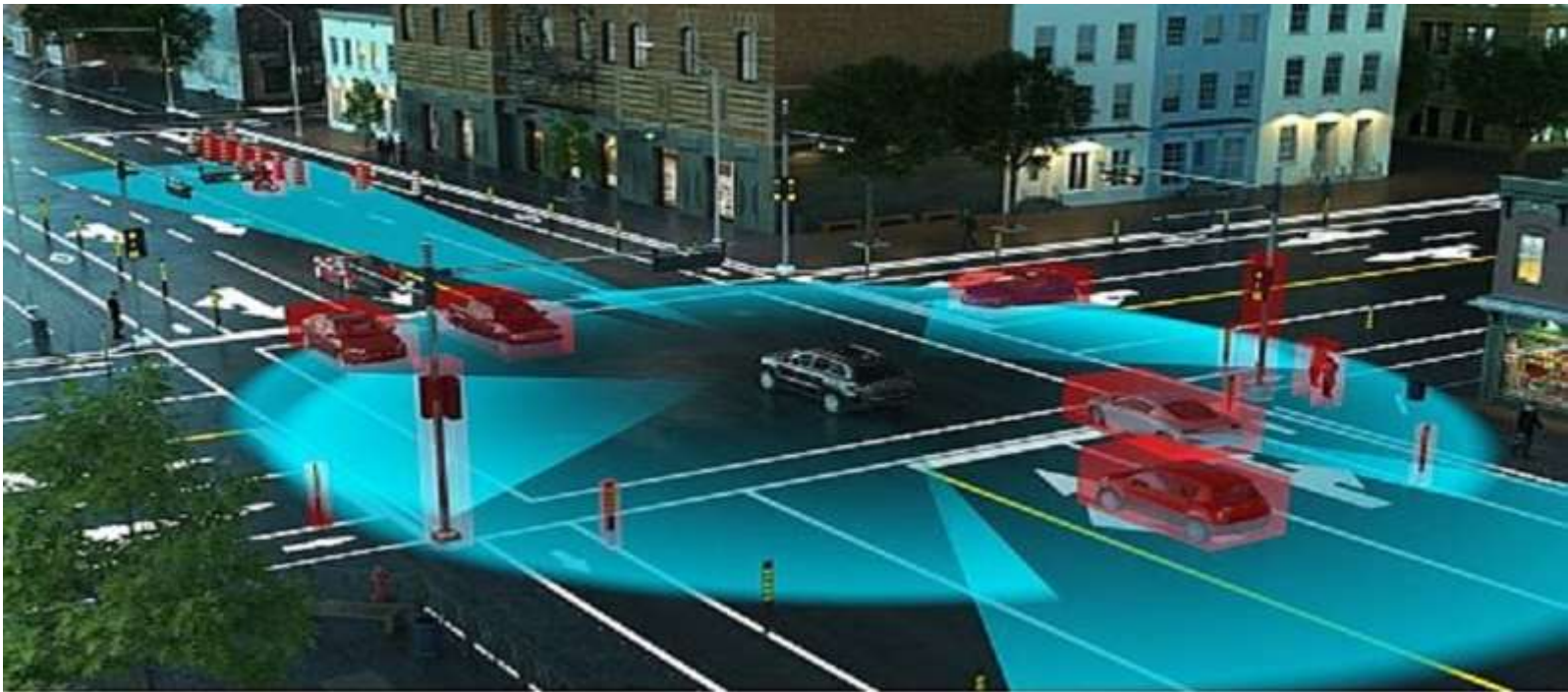
MSE (Mean Square Error)

RMSE (Root Mean Square Error)

R2 Score

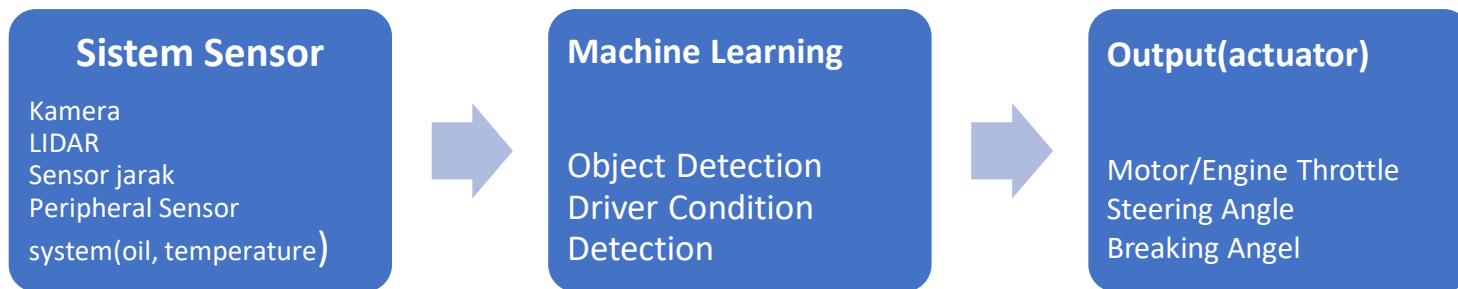
# Studi Kasus (Autonomous Vehicle)

Bagaimana menyediakan kendaraan tanpa pengemudi yang aman, ekonomis, dan praktis ?



<https://www.electronicproducts.com/autonomous-vehicles-and-the-truth-about-lidar-sensors/#>

# System Flow Autonomous Vehicle



Pada system ini dibutuhkan system sensor sebagai indra untuk mendeteksi lingkungan disekitar kendaraan.

Machine learning dibutuhkan untuk menangkap hasil pengindraan sensor dan menentukan actuator untuk bekerja.

# Komponen-Komponen Pendukung

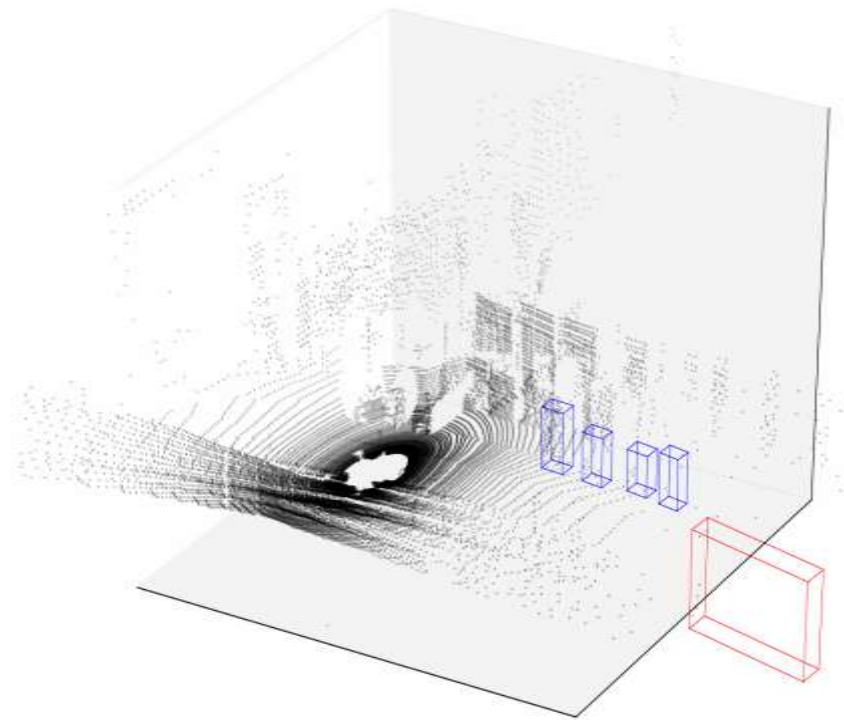
- Deteksi dan Klasifikasi Objek
- Pengenalan Pengemudi
- Sensor Fusion
- Aktuator
- Privacy



# Deteksi dan Klasifikasi Objek

Pembelajaran mesin digunakan untuk membantu decision making bagi pengemudi, seperti persepsi dan pemahaman visual disekitar kendaraan.

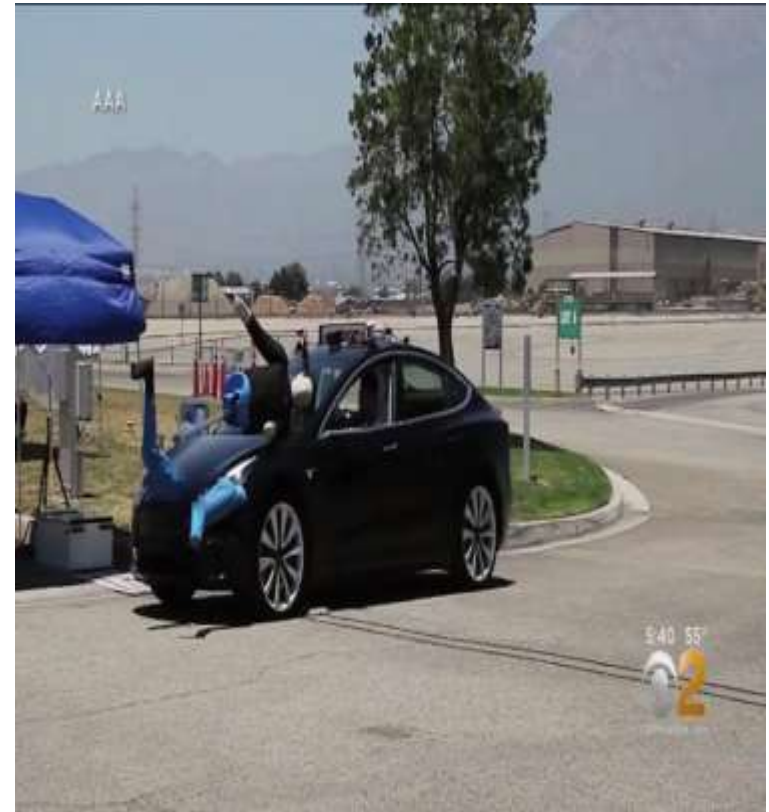
Hal ini melibatkan penggunaan system berbasis kamera untuk mendeteksi dan mengklasifikasikan objek, tetapi ada juga yang mengembangkan menggunakan LiDAR dan radar



<https://navoshta.com/kitti-lidar/>

# Deteksi dan Klasifikasi Objek (problems)

- Salah satu masalah terbesar dalam problem mobil otonom adalah salah klasifikasi objek atau keterlambatan dalam klasifikasi objek dan memberikan reaksi.
- Data yang dikumpulkan oleh berbagai sensor kendaraan dikumpulkan dan kemudian diinterpretasikan oleh sistem kendaraan



<https://newyork.cbslocal.com/2019/10/03/aaa-automatic-braking-pedestrians/>



# Deteksi dan Klasifikasi Objek (possible solution)

Melalui pelatihan model ML yang lebih baik dan data yang lebih umum, sistem dapat meningkatkan persepsi dan mengidentifikasi objek dengan akurasi yang lebih baik.

Melatih sistem dengan memberikan input yang lebih bervariasi pada parameter utama untuk dapat digunakan membuat keputusan.

# Pengenalan Pengemudi (problems)

Machine learning dapat mengenali pola, sehingga dapat digunakan di dalam kendaraan untuk memantau pengemudi.

Misalnya, pengenalan wajah dapat digunakan untuk mengidentifikasi pengemudi dan memverifikasi apakah dia memiliki hak tertentu, misalnya izin untuk menyalakan mobil, yang dapat membantu mencegah penggunaan dan pencurian.

Pengenalan kondisi pengemudi (mengantuk atau siap berkendara)



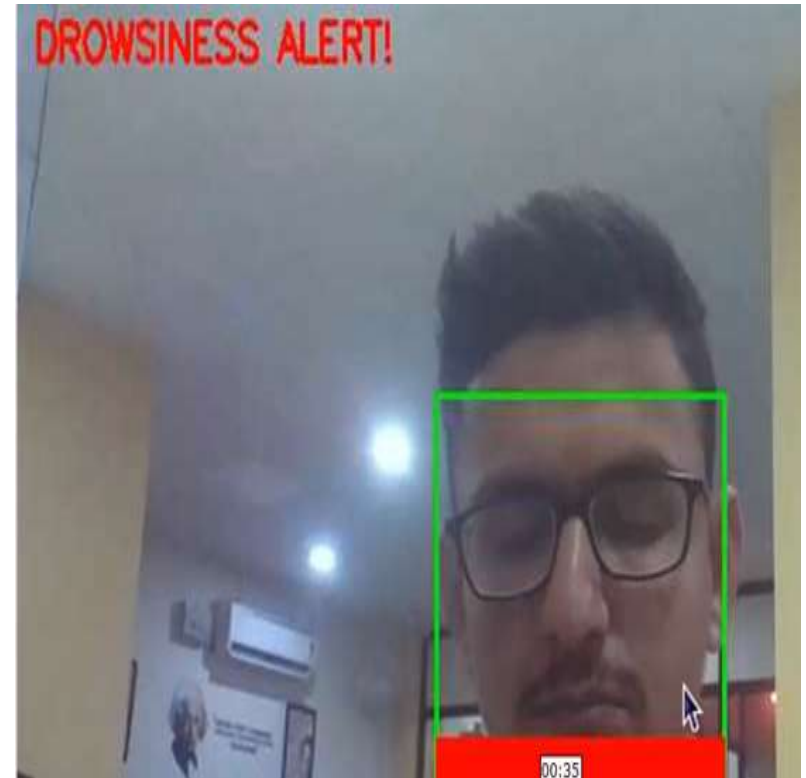
<https://www.openpr.com/news/679437/a-growing-driver-drowsiness-detection-system-market-2017-analysis-and-forecast-2021.html>

# Pengenalan Pengemudi (possible solutions)

Pengenalan wajah (face recognition) system perlu diimplementasikan untuk mengenali wajah pengemudi.

Selain itu untuk mendeteksi tingkat kelelahan kita memerlukan definisi dari face landmarks untuk membantu menentukan status kelelahan dari pengemudi.

Dengan menggunakan fitur-fitur yang berada pada face landmarks, kita bisa melihat lebih detail kondisi mata, kondisi mulut, hidung, dan dahi dari pengemudi.



<https://circuitdigest.com/microcontroller-projects/driver-drowsiness-detector-using-raspberry-pi-and-opencv>

# Privacy (problems)

Dengan semua data yang diambil dari sensor, muncul kekhawatiran yang dalam hal pengumpulan data . Hal ini perlu diatur agar diproses dilakukan secara legal dan benar

Dibutuhkan aturan yang meregulasi penggunaan data khusus untuk kendaraan otonom



<https://www.bosch-mobility-solutions.com/en/solutions/interior/interior-monitoring-systems/>