

Tratamento de dados

Como podemos tratar os dados dos dispositivos.

15 de outubro de 2020

Sumário

Tratamento
de dados

Tratamentos

Filtros

Aprendizado
de Máquina

1 Tratamentos

2 Filtros

3 Aprendizado de Máquina

Sumário

Tratamento
de dados

Tratamentos

Filtros

Aprendizado
de Máquina

1 Tratamentos

2 Filtros

3 Aprendizado de Máquina

Tratamentos de dados

Tratamento
de dados

Tratamentos

Filtros

Aprendizado
de Máquina

Tipos de tratamento

- Filtros
- Aprendizado
- Validação

Porque tratar os dados?

Tratamento
de dados

Tratamentos

Filtros

Aprendizado
de Máquina

Big Data Vs:

- Volume
- Velocidade
- Variedade
- Variabilidade
- Veracidade
- Validade
- Vulnerabilidade
- Volatilidade
- Visualização
- Valor

Sumário

Tratamento
de dados

Tratamentos

Filtros

Aprendizado
de Máquina

1 Tratamentos

2 Filtros

3 Aprendizado de Máquina

Filtros

Tratamento
de dados

Tratamentos

Filtros

Aprendizado
de Máquina

Tipos de filtros

- Analógicos
- Digitais
- Contextuais

Filtros

Tratamento
de dados

Tratamentos

Filtros

Aprendizado
de Máquina

Filtros comuns

- Passa-alta (High Pass Filter)
- Passa-baixa (Low Pass Filter)
- Passa-banda (Band Pass Filter)

Fusão de sensores

Tratamento
de dados

Tratamentos

Filtros

Aprendizado
de Máquina

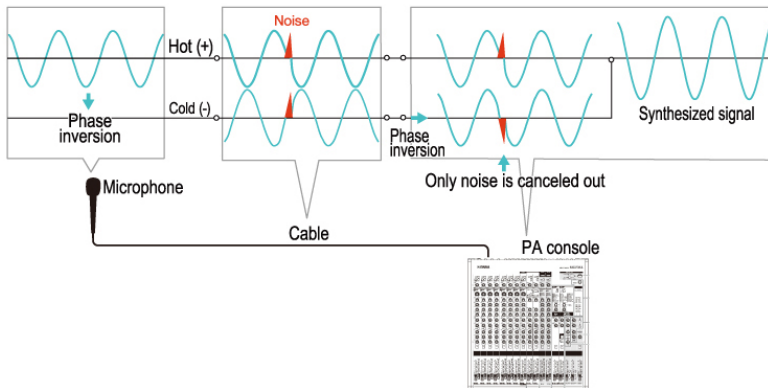
Utilidades

- Remover ruídos
- Decisões mais inteligentes
- Inferir informações

Fusão de sensores

Tratamento
de dados

Duplicação de sinal



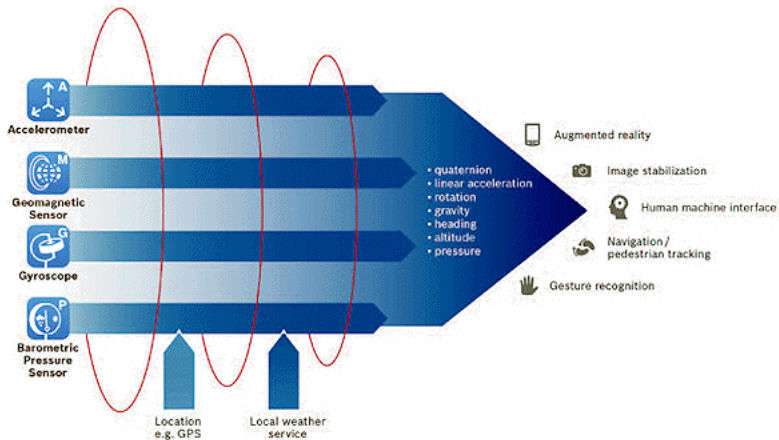
Tratamentos

Filtros

Aprendizado
de Máquina

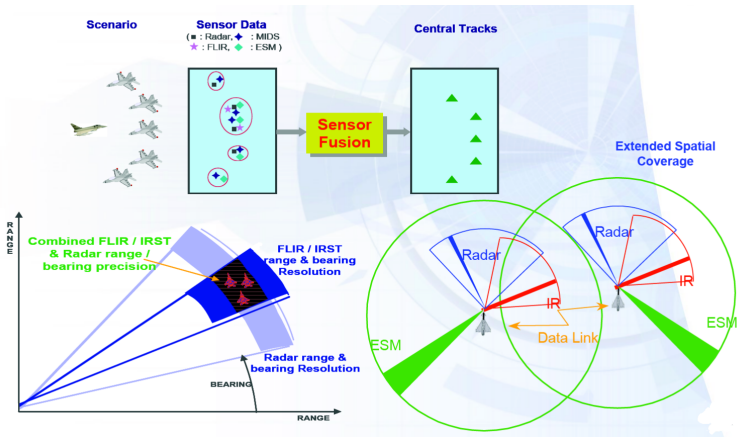
Fusão de sensores

Sensores magnéticos



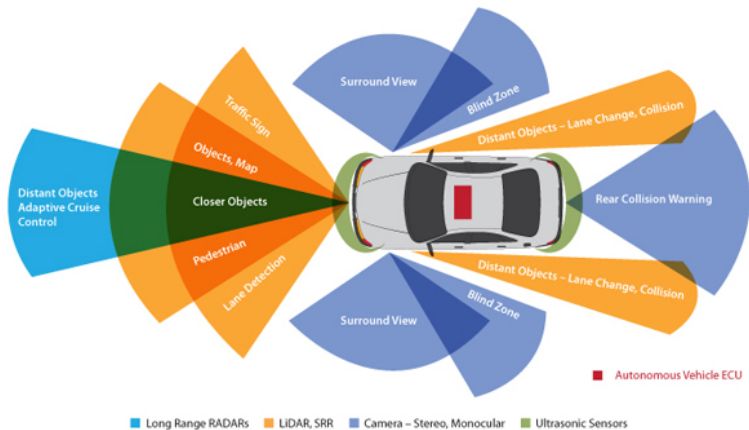
Fusão de sensores

Exemplo com radares



Fusão de sensores

Carro autônomo



Tratamento
de dados

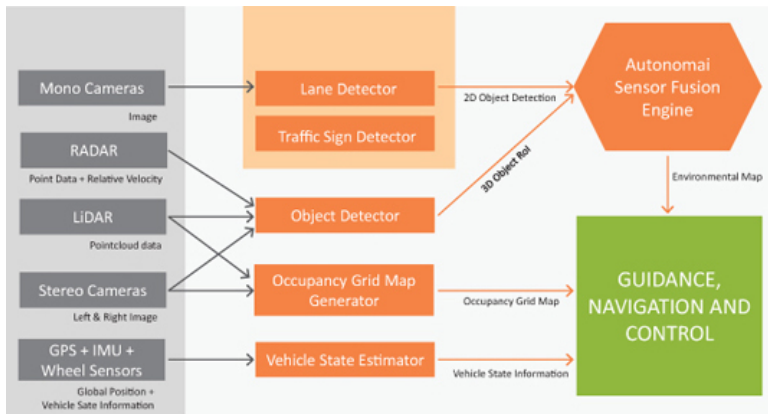
Tratamentos

Filtros

Aprendizado
de Máquina

Fusão de sensores

Carro autônomo



Sumário

Tratamento
de dados

Tratamentos
Filtros

Aprendizado
de Máquina

1 Tratamentos

2 Filtros

3 Aprendizado de Máquina

Aprendizado de máquina

Tratamento
de dados

Tratamentos

Filtros

Aprendizado
de Máquina

Por que?!

- Muitos dados
- Segurança para informações e sistemas
- Aumentar poder computacional
- Consumo eficiente de recursos e energia
- Crescimento constante de algoritmos e teorias

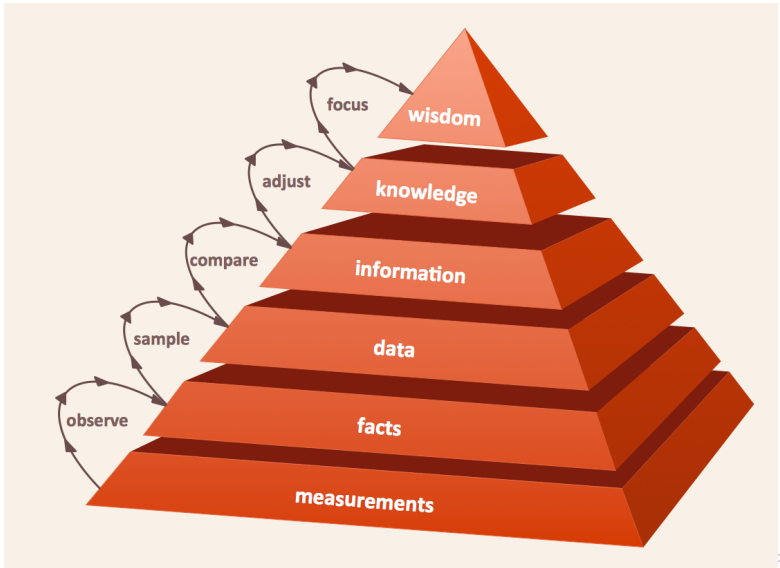
Aprendizado de máquina

Tratamento
de dados

Tratamentos

Filtros

Aprendizado
de Máquina



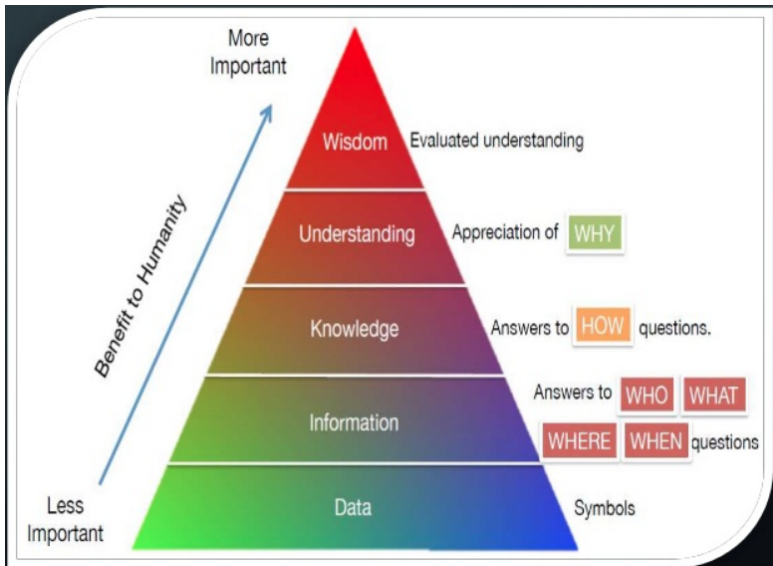
Aprendizado de máquina

Tratamento
de dados

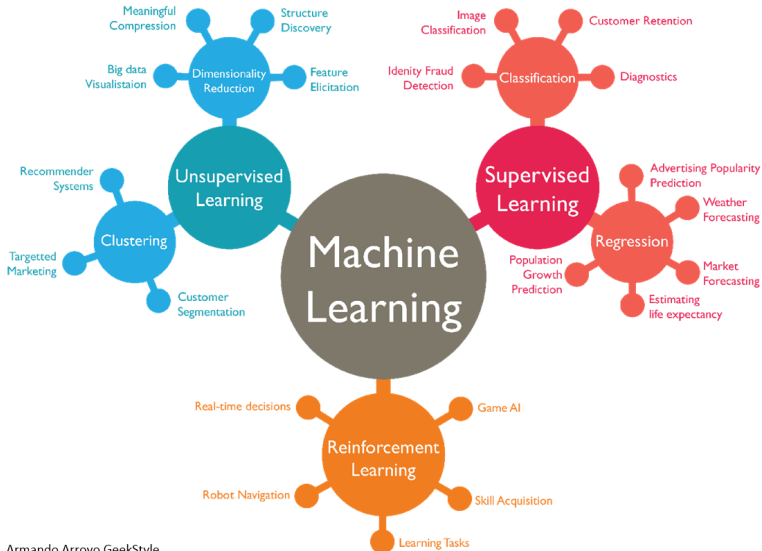
Tratamentos

Filtros

Aprendizado
de Máquina



Aprendizado de máquina



Aprendizado de máquina

Tratamento
de dados

Tratamentos

Filtros

Aprendizado
de Máquina

Algoritmos

■ Classificação

- K-Nearest Neighbors (KNN), Naive Bayes, Support Vector Machine (SVM)

■ Regressão

- Linear Regression, Support Vector Regression, Random Forests, Bagging

■ Clusterização

- K-Means, Density-Based Spatial Clustering of Applications with Noise

■ Extração de características

- Principal Component Analysis (PCA), Canonical Correlation Analysis, Feed Forward Neural Network

■ Detecção de anomalias

- One-class Support Vector Machines

Algoritmos

Tratamento
de dados

Tratamentos

Filtros

Aprendizado
de Máquina

K-Nearest Neighbors (KNN)

- Utilizado para classificar um novo dado desconhecido
- Utiliza um conjunto de dados de treinamento
- Para encontrar os vizinhos mais próximos utiliza:
 - Euclidean distance
 - L_{∞} norm
 - Ângulo
 - Mahalanobis distance
 - Hamming distance
- Requer armazenar o conjunto de dados de treinamento
- Exemplo de uso:
 - Classificação de regularidades em padrão de navegação

Algoritmos

Tratamento
de dados

Tratamentos

Filtros

Aprendizado
de Máquina

Naive Bayes

- Classificador probabilístico
- Utilizado para classificar um novo dado desconhecido
- Utiliza o Teorema de Bayes
- Considera ingenuamente a independência entre os atributos do novo dado
- Requer poucos dados para treinamento
- Trabalha com dados multi-dimensionais
- Rápido e escalável
- Exemplos de uso:
 - Categorização de texto, diagnóstico médico automático, confiança de produto da agricultura

Algoritmos

Tratamento
de dados

Tratamentos

Filtros

Aprendizado
de Máquina

Support Vector Machine (SVM)

- Classificador não-probabilístico e binário
- Busca encontrar o hiperplano que separa classes de dados de treinamento
- Um novo dado é classificado de acordo com sua posição em relação ao hiperplano
- Exemplos de uso:
 - Classificação de imagens, dados ambientais

Plataformas online

Tratamento
de dados

Tratamentos

Filtros

Aprendizado
de Máquina

Ferramentas online

- Microsoft Azure IoT Edge
- Amazon AWS IoT Greengrass ML Inference
- IBM Watson IoT Platform
- Google Cloud IoT
- Siemens IoT Mindsphere

Tratamento
de dados

Tratamentos

Filtros

Aprendizado
de Máquina

Tratamento de dados

Como podemos tratar os dados dos dispositivos.

15 de outubro de 2020