

“Supervised learning final project”

Diego Cubides,Larry castro, Tomas Mendez

2023-05-29

Abstract

The present report is based on the implementation of machine learning to predict three types of environments using the KNN algorithm, decision tree, random forest, and logistic regression. Data was collected from a DHT11 sensor, which measures humidity and temperature, an RGB color sensor TCS3200D TCS230, and a PT100 probe. RStudio was used to train the models with a dataset, and then the prediction was implemented for new data.

For data acquisition and communication, a robot controlled by an Arduino was used, communicating via serial communication with the HC-05 Bluetooth module. This allowed the robot to move easily between different environments and gather data characterizing each one.

Keywords: machine learning, predict, KNN algorithm, decision tree, random forest, logistic regression, DHT11 sensor, humidity, temperature, RGB color sensor TCS3200D TCS230, PT100 probe, RStudio, model training, prediction of new data, data acquisition, serial communication, Arduino, HC-05 Bluetooth module, robot, environments.

Resumen

El presente informe se centra en la aplicación del aprendizaje automático para predecir tres tipos de ambientes utilizando los algoritmos KNN, árbol de decisiones, bosque aleatorio y regresión logística. Se recopilaron datos de un sensor DHT11 para medir la humedad y la temperatura, un sensor de color RGB TCS3200D TCS230 y una sonda PT100. Se utilizó RStudio para entrenar los modelos utilizando un conjunto de datos y luego se implementó la predicción para nuevos datos.

Para la adquisición de datos y la comunicación, se utilizó un robot controlado por un Arduino, que se comunicaba a través de la comunicación serial con el módulo de Bluetooth HC-05. Esto permitió que el robot se moviera fácilmente entre diferentes ambientes y recopilara los datos que caracterizaban cada uno de ellos.

Palabras clave: aprendizaje de máquina, predecir, algoritmo Knn, árbol de decisiones, bosque aleatorio, regresión logística, sensor DHT11, humedad, temperatura, sensor de color RGB TCS3200D TCS230, sonda PT100, RStudio, entrenamiento de modelos, predicción de nuevos datos, adquisición de datos, comunicación serial, Arduino, módulo de Bluetooth HC-05, robot, ambientes.

Introducción

El aprendizaje de máquina y sus modelos, como KNN, árbol de decisiones, bosque aleatorio y regresión logística, son vitales en la actualidad. Estos modelos permiten resolver problemas complejos y tomar decisiones precisas. KNN clasifica ejemplos según su proximidad, mientras que el árbol de decisiones utiliza reglas jerárquicas. El bosque aleatorio combina múltiples árboles para mayor precisión, y la regresión logística predice probabilidades binarias. Estos modelos analizan grandes volúmenes de datos, revelando patrones ocultos y haciendo predicciones en diversos campos, desde medicina hasta finanzas. Su importancia radica en mejorar la toma de decisiones y generar conocimientos valiosos.

Materiales

- Sensor de color RGB TCS3200D TCS230
- Sensor DHT11
- Sonda PT100
- Modulo puente h l298n
- Carro a control remoto
- Bluetooth hc-05
- Arduino UNO
- Dispositivo Android

Software

- RStudio
- Arduino
- App inventor
- Excel
- PLX-DAQ

Procedimiento

adquisision de datos

Se desarrolló un sistema de control remoto para un carro mediante la tecnología Bluetooth, el cual posee la capacidad de moverse a diferentes velocidades y en múltiples direcciones. Para lograr esto, se creó una aplicación utilizando App Inventor, que permite el control preciso del movimiento del carro y la transmisión de datos en tiempo real. La app facilita la interacción con el carro y la recepción de los datos capturados en cada instante.



Figure 1: App

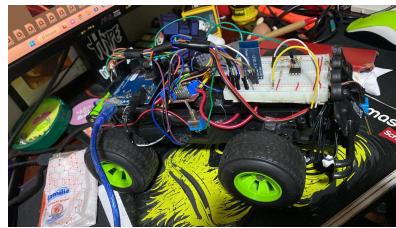


Figure 2: Robot

Para la adquisición de datos en tiempo real, se empleó el programa PLX-DAQ, el cual facilita la comunicación entre los datos provenientes del Arduino y Microsoft Excel. Esta integración permitió abrir los datos en RStudio como un archivo CSV, lo que posibilitó su posterior análisis y procesamiento.

#	A	B	C	D	E	F	G	H	I	J	K	L
1	BOARD	VTHRE	ACIN	TEMP	HUMI	PT100	AMBIENTE					
2	23	27	26	60	17	508	deserto					
3	25	27	27	57	60	509	deserto					
4	25	27	27	57	60	510	deserto					
5	47	55	72	38	38	525	topico					
6	34	31	45	61	33	526	topico					
7	24	24	21	21	24	527	inverno					
8	24	25	21	21	22	528	inverno					
9	24	25	21	21	22	529	inverno					
10	19	25	42	35	35	530	deserto					
11	4	4	4	88	88	531	deserto					
12	30	4	4	88	88	532	deserto					
13	25	25	25	25	25	533	deserto					
14	30	30	30	30	30	534	deserto					
15	30	30	30	30	30	535	deserto					
16	30	30	30	30	30	536	deserto					
17	30	30	30	30	30	537	deserto					
18	30	30	30	30	30	538	deserto					
19	30	30	30	30	30	539	deserto					
20	30	30	30	30	30	540	deserto					
21	30	30	30	30	30	541	deserto					
22	30	30	30	30	30	542	deserto					
23	30	30	30	30	30	543	deserto					
24	30	30	30	30	30	544	deserto					
25	30	30	30	30	30	545	deserto					
26	30	30	30	30	30	546	deserto					
27	30	30	30	30	30	547	deserto					
28	30	30	30	30	30	548	deserto					
29	30	30	30	30	30	549	deserto					
30	30	30	30	30	30	550	deserto					

Figure 3: PLX-DQ

Entrenamiento e implementacion de los modelos

Se desarrolló un sistema de control remoto para un carro mediante la tecnología Bluetooth, el cual posee la capacidad de moverse a diferentes velocidades y en múltiples direcciones. Para lograr esto, se creó una aplicación utilizando App Inventor, que permite el control preciso del movimiento del carro y la transmisión de datos en tiempo real. La app facilita la interacción con el carro y la recepción de los datos capturados en cada instante.