



ESCUELA DE INGENIERÍA DE  
FUENLABRADA

Grado en Ingeniería de Sistemas de Telecomunicación

## Trabajo Fin de Grado

**Posicionamiento en interiores utilizando luz visible y un  
fotodiodo a través de algoritmos de regresión**

por

Mariano Jiménez Ruiz

Tutora: Elena Aparicio Esteve

Co-tutor: David Casillas Pérez

Año Académico 2023/2024







# Agradecimientos

Tras todo el esfuerzo realizado me gustaría dar las gracias tanto a mi tutora, Elena Aparicio Esteve, como a mi co-tutor, David Casillas Pérez, por su dedicación y trabajo ayudándome en todo lo posible para que yo pudiera realizar este TFG.

También me gustaría agradecer a mis compañeros y amigos del grado con los que he compartido muchos momentos, Aitor Clemente Caballero, Juan Manuel Ruiz Casado, Veselin Todorov Hristov, Rufino García Sánchez y Victor Rincón Yepes.

Por último, agradecer a mi familia, en concreto a mis padres, Mariano Jiménez Díaz y María Milagros Ruiz Martín, así como a mi hermana Mónica Jiménez Ruiz, por su apoyo incondicional desde el primer momento y su sacrificio para que yo pudiera realizar y acabar este grado.

# Abstract

Humanity always has pursued knowledge with the aim of understanding how the world works and, in that way improve people's lives. As society made progress, new challenges were found out to apply them to daily life. This is how the quest to obtain the location both people and objects. After some years of research, GPS was developed and implemented. But this technology is not perfect because it can be applied only outdoors due to their own performance. When the emitted signals try to pass through certain materials, attenuation, reflection and refraction affect them reducing their intensity or even signal loss, causing system failure. In the face of lacking a definitive method for indoors, different investigations have been carried out in order to solve this problem. These investigations have been possible thanks to the development of ML in recent years. Thus by using jointly ML and devices such a photodiode, which received light intensity from LEDs, and after that an subsequent analysis and research of the data, a complex system, which allows indoors location tracking, is created. So this is the aim of this TFG.

The current work focuses on proving that using different regression algorithms can get the position of a receiver indoors. The accuracy of various regressors in estimating the receiver's coordinates from data collected is evaluated. For that purpose, an already existing database, which has been created in another investigation is used. This is implemented through an experiment based on putting four LEDs in the ceiling at a height of 1,284m from a floor plan with an area of 3x3m, in which a receiver was placed in different locations and received light intensity was measured for that point depending on the LED. Thus, the receiver contained a photo diode. In addition, each LED beamed light with different frequencies. So a database was built with  $x$  and  $y$  coordinates of the receiver and also another one with the intensity for each position and LED.

Firstly, I have carried out an analysis of the data, which consists of  $x$  and  $y$  coordinates of the receiver and the intensity at each position for each LED. Later, two kind of experiments are executed. On the one hand, the type 1 experiment, in which thanks to the database the coordinates of the positions are estimated. On the other hand, the type 2, where what is estimated is the distance among the receiver and the LEDs and after

all calculating the coordinates through the intersection of circumferences. These experiments are repeated for different regressors, like linear regressor, Support Vector Regression (SVR) linear regressor, SVR polynomial regressor and SVR rbf regressor. In this way, the database is separated into two groups, one for training and the another one for testing. Besides, this process is carried out with different combinations of LEDs. In other words, as there are four LEDs it is interesting to compare the results using one, two, three or four of the LEDs. Furthermore, it has been experimented 100 times in order to ensure that both the process and the data are right implementing the cross validation method.

After that, the data are represented in various graphs with the purpose of a more visually analysis and they are gathered in tables to have the information well organised. By doing so, the results show that the type 2 experiment is more accurate and therefore get low values of all the errors calculated (RMSE, MAE and MAER). It is interesting to point out the fact that the best combination of LEDs according to the results is the one with the four LEDs. Lastly, just mentioning the SVR RBF regressor shows the best performance.

Thanks to this work, it is demonstrated that it is possible to get the position indoors with regression algorithms. This fact can suppose a breakthrough in many processes either logistics sector or business area allowing to offer more personal experiences. All of this, without expending a great amount of money. Moreover, this fact enables the continuity of investigations in this field with other regression algorithms or varying the number of LEDs.





# Resumen

La humanidad siempre ha perseguido el conocimiento con el fin de entender como funciona el mundo y, así, mejorar la vida de las personas. A medida que avanzaba la sociedad, se encontraban nuevas cuestiones a resolver que se pudieran aplicar a la vida diaria. Es así como surge la búsqueda para obtener la localización tanto de personas como de objetos. Tras años de estudio se consiguió desarrollar e implementar la tecnología de GPS. Pero ésta tiene un inconveniente y es que solo se puede aplicar a exteriores debido al propio funcionamiento del sistema. Ya que las señales emitidas al intentar atravesar ciertos materiales pueden sufrir los efectos de atenuación, refracción o reflexión pudiendo reducir la intensidad de la señal o incluso la pérdida total de la misma, lo que produce un fallo del sistema. Ante la falta de un método definitivo para interiores, se han llevado a cabo diferentes investigaciones para poder solventar esta problemática. Estas investigaciones se han producido gracias al desarrollo del ML en los últimos años. De tal modo, que la aplicación conjunta del ML junto con dispositivos como son fotodiodos, que reciben la intensidad de la luz de LEDs, y un posterior análisis e investigación de estos datos, se puede producir un sistema más complejo que permita obtener la ubicación en interiores. Así, surge la finalidad de este TFG.

Este TFG aborda la aplicación de distintos algoritmos de regresión para obtener la posición de un receptor en entornos interiores. Se evaluó la precisión de distintos regresores en la estimación de las coordenadas del receptor a partir de los datos recogidos en una base de datos. Para ello, se ha usado una base de datos ya existente y creada en otra investigación. Ésta se ha obtenido mediante un experimento en el que se colocaban cuatro LEDs en el techo a una altura de 1,284m de un plano base con una superficie de 3x3m, en el que se colocaba un receptor en distintos lugares del plano y se medía la intensidad de la luz recibida en ese punto por cada LED. Además, cada LED emitía a una frecuencia distinta.

En primer lugar, se ha llevado a cabo un análisis de la base de datos, que consta de las coordenadas  $x$  e  $y$  del receptor y la intensidad en cada posición para cada LED. Después, se han realizado dos tipos de experimentos. Por un lado, el experimento tipo

1, estima directamente las coordenadas de las posiciones del receptor a partir de los datos de la intensidad de la luz. Por otro lado, el tipo 2 en el que se calcula la distancia entre la posición del receptor y las posiciones de cada LED, con el objetivo de luego determinar las coordenadas del receptor mediante la intersección de circunferencias. Estos dos experimentos se han repetido para distintos regresores, los cuales son: regresor lineal, regresor SVR lineal, regresor SVR polinomial y regresor SVR RBF. De tal forma, que se han separado los datos en dos grupos uno de entrenamiento y uno de test. Además, esta secuencia se ha realizado con las distintas combinaciones de LEDs. Es decir, ya que había cuatro LEDs, se quería ver como variaban los resultados si en vez de disponer de datos de cuatro LEDs, solo había datos de uno, dos, tres o de los cuatro. Igualmente, para asegurarse que los datos y el método eran correctos, todo este proceso se ha realizado un total de 100 veces aplicando así una técnica de validación cruzada.

Tras este procedimiento, se han representado los datos en distintas gráficas y se han recogido en diferentes tablas para un análisis detallado. Así, se obtiene que el experimento tipo 2 es más preciso y, por tanto, obtiene menores valores en todos los errores que se han calculado (RMSE, MAE y MAER). También cabe destacar que la combinación con la que los resultados estimados se acercan más a los reales es la combinación de los cuatro LEDs. Por último, mencionar el regresor que mejor desempeño muestra a la hora de estimar estas posiciones que es el regresor SVR RBF.

Gracias a este TFG, se demuestra que con distintos algoritmos de regresión se puede obtener la ubicación precisa (centimétrica) en interiores. Lo que puede ayudar en gran medida al avance de muchos procesos ya sea en el sector de la logística o en el sector comercial permitiendo ofrecer experiencias más personales. Todo ello, sin necesidad de una gran inversión económica. Además, esto permite la posibilidad de futuras investigaciones en este campo con el uso de otros algoritmos de regresión o la variación en el número de LEDs.

# Índice general

<b>Agradecimientos</b>	<b>I</b>
<b>Abstract</b>	<b>II</b>
<b>Resumen</b>	<b>V</b>
<b>Listado of Figuras</b>	<b>XIV</b>
<b>Listado of Tablas</b>	<b>XX</b>
<b>Acrónimos</b>	<b>XXII</b>
<b>1. Introducción</b>	<b>1</b>
1.1. Motivación . . . . .	3
1.2. Objetivos . . . . .	3
1.3. Estructura de la memoria . . . . .	4
<b>2. Estado del arte</b>	<b>6</b>
2.1. Introducción . . . . .	6
2.2. Trabajos de VLP . . . . .	7
<b>3. Estudio teórico</b>	<b>16</b>
3.1. Machine Learning . . . . .	16
3.2. Validación cruzada . . . . .	19
3.3. Regresores . . . . .	22
3.3.1. Regresión Lineal . . . . .	22
3.3.2. SVR . . . . .	23
3.4. Mínimos cuadrados . . . . .	25

<b>4. Metodología</b>	<b>27</b>
4.1. Introducción . . . . .	27
4.2. Base de datos . . . . .	28
4.3. Análisis de la base de datos . . . . .	29
4.4. Estimación de la posición del receptor . . . . .	35
4.5. Métricas . . . . .	36
4.5.1. RMSE . . . . .	37
4.5.2. MAE . . . . .	37
4.5.3. MAER . . . . .	38
<b>5. Resultados</b>	<b>39</b>
5.1. Introducción . . . . .	39
5.2. Resultados . . . . .	39
5.2.1. Regresor lineal . . . . .	39
5.2.2. Regresor SVR Lineal . . . . .	45
5.2.3. Regresor SVR polinomial . . . . .	49
5.2.4. Regresor SVR RBF . . . . .	53
5.3. Comparativa . . . . .	58
<b>6. Conclusiones y Líneas futuras</b>	<b>60</b>
6.1. Conclusiones . . . . .	60
6.2. Impacto . . . . .	61
6.2.1. Impacto social . . . . .	61
6.2.2. Impacto medioambiental . . . . .	62
6.2.3. Impacto económico . . . . .	62
6.3. Lecciones aprendidas . . . . .	64
6.4. Líneas futuras . . . . .	64
<b>A. Apéndice</b>	<b>66</b>
A.1. Capítulo 4 - Metodología . . . . .	66
A.2. Capítulo 5 - Resultados . . . . .	66

A.2.1. Tablas del Regresor lineal . . . . .	66
A.2.2. Gráficas del Regresor lineal . . . . .	70
A.2.3. Tablas del Regresor SVR lineal . . . . .	74
A.2.4. Gráficas del Regresor SVR lineal . . . . .	78
A.2.5. Tablas del Regresor SVR polinomial . . . . .	82
A.2.6. Gráficas del Regresor SVR polinomial . . . . .	86
A.2.7. Tablas del Regresor SVR rbf . . . . .	90
A.2.8. Gráficas del Regresor SVR rbf . . . . .	94

<b>Bibliografía</b>	<b>104</b>
---------------------	------------

# Listado of Figuras

1.1.	Escenario con la implantación de sensores y luces LED. . . . .	2
2.1.	Diseño del receptor donde se agrupan cuatro fotodiodos. . . . .	9
2.2.	Montaje tecnología HTC Vive. . . . .	10
2.3.	Diseño del cuadrante de celdas solares. . . . .	10
2.4.	Funcionamiento del software <i>DIALux</i> . . . . .	11
2.5.	Segmentación de una imagen. . . . .	12
2.6.	Segmentación de imágenes obtenidas por una cámara. . . . .	13
3.1.	Ejemplo de algoritmo de clasificación. . . . .	18
3.2.	Ejemplos de agrupación de datos en clusters. . . . .	19
3.3.	Diagrama resumen de la técnica de <i>Cross-validation</i> . . . . .	20
3.4.	Diagrama de la técnica de <i>Single hold-out random subsampling</i> . . . . .	20
3.5.	Diagrama de la técnica de <i>k-fold cross-validation</i> . . . . .	21
3.6.	Diagrama de la técnica de <i>leave-one-out cross-validation</i> . . . . .	21
3.7.	<i>Support Vector Regression</i> - Hiperplano . . . . .	23
3.8.	Transformación de los datos usando un kernel. . . . .	24
4.1.	Esquema descriptivo de las pasos seguidos durante el TFG. . . . .	27
4.2.	Escenario usado en la toma de datos. . . . .	28
4.3.	Escenario con la posición de cada LED. . . . .	29
4.4.	Distribución de la base de datos sobre el plano XY. . . . .	30
4.5.	Intensidad recibida de cada LED en cada posición separadas por LEDs por colores. . . . .	31
4.6.	Intensidad recibida de cada LED en cada posición separadas por LEDs en distintas gráficas. . . . .	31

4.7. Histograma Intensidad-Frecuencia separados por colores por LEDs. . . . .	31
4.8. Histograma Intensidad-Frecuencia y curva de aproximación separados por gráficas por LEDs. . . . .	33
4.9. Gráfico de como se cácula la distancia entre un LED y una posición. . . . .	34
4.10. Relación Intensidad-Distancia separadas por gráficas por LEDs. . . . .	34
5.1. Box-plot del RMSE para el regresor lineal. . . . .	40
5.2. Relación Intensidad-Distancia de los datos reales y los estimados para el LED 1 con el experimento tipo 1. . . . .	41
5.3. Relación Intensidad-Distancia de los datos reales y los estimados para el LED 1 con el experimento tipo 2. . . . .	41
5.4. CDF del RMSE de las coordenadas $(x,y)$ obtenidas para el experimento tipo 1 (coord.) y el experimento tipo 2 (dist.). . . . .	42
5.5. CDF del RMSE de las distancias obtenidas para el experimento tipo 2. . . . .	42
5.6. Box-plot del RMSE para el regresor lineal. . . . .	43
5.7. Distribución de los datos reales y los estimados sobre el plano para el regresor SVR lineal. . . . .	45
5.8. Relación Intensidad-Distancia de los datos reales y los estimados para el LED 1 con el experimento tipo 1. . . . .	46
5.9. Relación Intensidad-Distancia de los datos reales y los estimados para el LED 1 con el experimento tipo 2. . . . .	46
5.10. CDF del RMSE de las coordenadas $(x,y)$ obtenidas para el experimento tipo 1 (coord.) y el experimento tipo 2 (dist.). . . . .	47
5.11. CDF del RMSE de las distancias obtenidas para el experimento tipo 2. . . . .	47
5.12. Boxplot del RMSE para el regresor SVR lineal. . . . .	48
5.13. Distribución de los datos reales y los estimados sobre el plano para el regresor SVR polinomial. . . . .	50
5.14. Relación Intensidad-Distancia de los datos reales y los estimados para el LED 1 con el experimento tipo 2. . . . .	50
5.15. Relación Intensidad-Distancia de los datos reales y los estimados para el LED 1 con el experimento tipo 2. . . . .	51
5.16. CDF del RMSE de las coordenadas $(x,y)$ obtenidas para el experimento tipo 1 (coord.) y el experimento tipo 2 (dist.). . . . .	51

5.17. CDF del RMSE de las distancias obtenidas para el experimento tipo 2. . .	52
5.18. Boxplot del RMSE para el regresor SVR polinomial. . . . .	52
5.19. Distribución de los datos reales y los estimados sobre el plano para el re- gresor SVR RBF. . . . .	54
5.20. Relación Intensidad-Distancia de los datos reales y los estimados para el LED 1 con el experimento tipo 1. . . . .	55
5.21. Relación Intensidad-Distancia de los datos reales y los estimados para el LED 1 con el experimento tipo 2. . . . .	55
5.22. CDF del RMSE de las coordenadas $(x,y)$ obtenidas para el experimento tipo 1 (coord.) y el experimento tipo 2 (dist.). . . . .	56
5.23. CDF del RMSE de las distancias obtenidas para el experimento tipo 2. . .	56
5.24. Box-plot de los RMSE para el regresor SVR RBF. . . . .	57
A.1. Histograma Intensidad-Frecuencia por colores para cada LED. . . . .	66
A.2. Relación Intensidad-Distancia de los datos reales y los estimados para el LED 2. . . . .	70
A.3. Relación Intensidad-Distancia de los datos reales y los estimados para el LED 3. . . . .	71
A.4. Relación Intensidad-Distancia de los datos reales y los estimados para el LED 4. . . . .	71
A.5. CDF del MAE para el experimento tipo 1. . . . .	71
A.6. CDF del MAER para el experimento tipo 1. . . . .	72
A.7. CDF del MAE para el experimento tipo 2. . . . .	72
A.8. CDF del MAER para el experimento tipo 2. . . . .	72
A.9. Relación Intensidad-Distancia de los datos reales y los estimados para el LED 2 para el experimento tipo 2. . . . .	73
A.10. Relación Intensidad-Distancia de los datos reales y los estimados para el LED 3 para el experimento tipo 2. . . . .	73
A.11. Relación Intensidad-Distancia de los datos reales y los estimados para el LED 4 para el experimento tipo 2. . . . .	73
A.12. Box-plot de los MAE y MAER para el regresor lineal. . . . .	74
A.13. Relación Intensidad-Distancia de los datos reales y los estimados para el LED 2. . . . .	78



A.14.Relación Intensidad-Distancia de los datos reales y los estimados para el LED 3. . . . .	78
A.15.Relación Intensidad-Distancia de los datos reales y los estimados para el LED 4. . . . .	79
A.16.CDF del MAE para el experimento tipo 1. . . . .	79
A.17.CDF del MAER para el experimento tipo 1. . . . .	79
A.18.CDF del MAE para el experimento tipo 2. . . . .	80
A.19.CDF del MAER para el experimento tipo 2. . . . .	80
A.20.Relación Intensidad-Distancia de los datos reales y los estimados para el LED 2 para el experimento tipo 2. . . . .	80
A.21.Relación Intensidad-Distancia de los datos reales y los estimados para el LED 3 para el experimento tipo 2. . . . .	81
A.22.Relación Intensidad-Distancia de los datos reales y los estimados para el LED 4 para el experimento tipo 2. . . . .	81
A.23.Box-plot de los MAE y MAER para el regresor SVR lineal. . . . .	82
A.24.Relación Intensidad-Distancia de los datos reales y los estimados para el LED 2. . . . .	86
A.25.Relación Intensidad-Distancia de los datos reales y los estimados para el LED 3. . . . .	86
A.26.Relación Intensidad-Distancia de los datos reales y los estimados para el LED 4. . . . .	87
A.27.CDF del MAE para el experimento tipo 2. . . . .	87
A.28.CDF del MAER para el experimento tipo 1. . . . .	87
A.29.CDF del MAE para el experimento tipo 2. . . . .	88
A.30.CDF del MAER para el experimento tipo 2. . . . .	88
A.31.Relación Intensidad-Distancia de los datos reales y los estimados para el LED 2 para el experimento tipo 2. . . . .	88
A.32.Relación Intensidad-Distancia de los datos reales y los estimados para el LED 3 para el experimento tipo 2. . . . .	89
A.33.Relación Intensidad-Distancia de los datos reales y los estimados para el LED 4 para el experimento tipo 2. . . . .	89
A.34.Box-plot de los MAE y MAER para el regresor SVR polinomial. . . . .	90

A.35.Relación Intensidad-Distancia de los datos reales y los estimados para el LED 2. . . . .	94
A.36.Relación Intensidad-Distancia de los datos reales y los estimados para el LED 3. . . . .	94
A.37.Relación Intensidad-Distancia de los datos reales y los estimados para el LED 4. . . . .	95
A.38.CDF del MAE para el experimento tipo 2. . . . .	95
A.39.CDF del MAER para el experimento tipo 1. . . . .	95
A.40.CDF del MAE para el experimento tipo 2. . . . .	96
A.41.CDF del MAER para el experimento tipo 2. . . . .	96
A.42.Relación Intensidad-Distancia de los datos reales y los estimados para el LED 2 para el experimento tipo 2. . . . .	96
A.43.Relación Intensidad-Distancia de los datos reales y los estimados para el LED 3 para el experimento tipo 2. . . . .	97
A.44.Relación Intensidad-Distancia de los datos reales y los estimados para el LED 4 para el experimento tipo 2. . . . .	97
A.45.Box-plot de los MAE y MAER para el regresor SVR RBF. . . . .	98

# Listado of Tablas

1.1. Listado de las aplicaciones nombradas en las que se puede aplicar la tecnología LED con el ML. . . . .	3
2.1. Resumen de los errores medios para LRML, KRRML con y sin SFDP. . . . .	9
2.2. Resumen de los trabajos analizados sobre VLP 1. . . . .	14
2.3. Resumen de los trabajos analizados sobre VLP 2. . . . .	15
5.1. Error cuadrático medio para el regresor lineal en el caso del experimento tipo 1 para cada combinación de LEDs. . . . .	44
5.2. Error cuadrático medio para el regresor lineal en el experimento tipo 2 para obtener unas coordenadas para cada combinación de LEDs. . . . .	44
5.3. Error cuadrático medio para el regresor lineal en el caso de estimar las distancias para cada combinación de LEDs. . . . .	44
5.4. Error cuadrático medio para el regresor svr lineal en el caso del experimento tipo 1 para cada combinación de LEDs. . . . .	49
5.5. Error cuadrático medio para el regresor svr lineal en el caso del experimento tipo 2 para cada combinación de LEDs. . . . .	49
5.6. Error cuadrático medio para el regresor svr lineal en el caso de estimar las distancias para cada combinación de LEDs. . . . .	49
5.7. Error cuadrático medio para el regresor svr polinomial en el caso del experimento tipo 1 para cada combinación de LEDs. . . . .	53
5.8. Error cuadrático medio para el regresor svr polinomial en el caso del experimento tipo 2 para cada combinación de LEDs. . . . .	53
5.9. Error cuadrático medio para el regresor svr polinomial en el caso de estimar las distancias para cada combinación de LEDs. . . . .	53
5.10. Error cuadrático medio para el regresor SVR rbf en el caso del experimento tipo 1 para cada combinación de LEDs. . . . .	57

5.11. Error cuadrático medio para el regresor SVR rbf en el caso del experimento 2 para cada combinación de LEDs. . . . .	58
5.12. Error cuadrático medio para el regresor SVR rbf en el caso de estimar las distancias para cada combinación de LEDs. . . . .	58
5.13. Errores cuadráticos medios para todos los regresores en el caso del experimento tipo 1 para la combinación de los 4 LEDs. . . . .	59
5.14. Errores cuadráticos medios para todos los regresores en el caso del experimento tipo 2 para la combinación de los 4 LEDs. . . . .	59
5.15. Errores cuadráticos medios para todos los regresores en el caso de estimar las distancias para la combinación de los 4 LEDs. . . . .	59
6.1. Metas específicas de los ODS que se adecuan a este TFG. . . . .	63
A.1. Error cuadrático medio para el regresor lineal en el caso del experimento tipo 1 para cada combinación de LEDs 1. . . . .	66
A.2. Error cuadrático medio para el regresor lineal en el caso del experimento tipo 1 para cada combinación de LEDs 2. . . . .	67
A.3. Error cuadrático medio para el regresor lineal en el caso del experimento tipo 2 para cada combinación de LEDs 1. . . . .	67
A.4. Error cuadrático medio para el regresor lineal en el caso del experimento tipo 2 para cada combinación de LEDs 2. . . . .	67
A.5. Error medio absoluto para el regresor lineal en el caso del experimento tipo 1 para cada combinación de LEDs 1. . . . .	67
A.6. Error medio absoluto para el regresor lineal en el caso del experimento tipo 1 para cada combinación de LEDs 2. . . . .	67
A.7. Error medio absoluto para el regresor lineal en el caso del experimento tipo 2 para cada combinación de LEDs 1. . . . .	68
A.8. Error medio absoluto para el regresor lineal en el caso del experimento tipo 2 para cada combinación de LEDs 2. . . . .	68
A.9. Error relativo medio para el regresor lineal en el caso del experimento tipo 1 para cada combinación de LEDs 1. . . . .	68
A.10. Error relativo medio para el regresor lineal en el caso del experimento tipo 1 para cada combinación de LEDs 2. . . . .	68
A.11. Error relativo medio para el regresor lineal en el caso del experimento tipo 2 para cada combinación de LEDs 1. . . . .	68

A.12. Error relativo medio para el regresor lineal en el caso del experimento tipo 2 para cada combinación de LEDs 2. . . . .	69
A.13. Error medio absoluto para el regresor lineal en el caso del experimento tipo 1 para cada combinación de LEDs 3. . . . .	69
A.14. Error medio absoluto para el regresor lineal en el caso del experimento tipo 2 para cada combinación de LEDs 3. . . . .	69
A.15. Error medio absoluto para el regresor lineal en el caso del experimento tipo 2 para cada combinación de LEDs. . . . .	69
A.16. Error relativo medio para el regresor lineal en el caso del experimento tipo 1 para cada combinación de LEDs 3. . . . .	69
A.17. Error relativo medio para el regresor lineal en el caso del experimento tipo 2 para cada combinación de LEDs 3. . . . .	70
A.18. Error relativo medio para el regresor lineal en el caso del experimento tipo 2 para cada combinación de LEDs. . . . .	70
A.19. Error cuadrático medio para el regresor SVR lineal en el caso del experimento tipo 1 para cada combinación de LEDs 1. . . . .	74
A.20. Error cuadrático medio para el regresor SVR lineal en el caso del experimento tipo 1 para cada combinación de LEDs 2. . . . .	74
A.21. Error cuadrático medio para el regresor SVR lineal en el caso del experimento tipo 2 para cada combinación de LEDs 1. . . . .	75
A.22. Error cuadrático medio para el regresor SVR lineal en el caso del experimento tipo 2 para cada combinación de LEDs 2. . . . .	75
A.23. Error medio absoluto para el regresor SVR lineal en el caso del experimento tipo 1 para cada combinación de LEDs 1. . . . .	75
A.24. Error medio absoluto para el regresor SVR lineal en el caso del experimento tipo 1 para cada combinación de LEDs 2. . . . .	75
A.25. Error medio absoluto para el regresor SVR lineal en el caso del experimento tipo 2 para cada combinación de LEDs 1. . . . .	75
A.26. Error medio absoluto para el regresor SVR lineal en el caso del experimento tipo 2 para cada combinación de LEDs 2. . . . .	76
A.27. Error relativo medio para el regresor SVR lineal en el caso del experimento tipo 1 para cada combinación de LEDs 1. . . . .	76
A.28. Error relativo medio para el regresor SVR lineal en el caso del experimento tipo 1 para cada combinación de LEDs 2. . . . .	76

A.29. Error relativo medio para el regresor SVR lineal en el caso del experimento tipo 1 para cada combinación de LEDs 1. . . . .	76
A.30. Error relativo medio para el regresor SVR lineal en el caso del experimento tipo 1 para cada combinación de LEDs 2. . . . .	76
A.31. Error medio absoluto para el regresor SVR lineal en el caso del experimento tipo 1 para cada combinación de LEDs 3. . . . .	77
A.32. Error medio absoluto para el regresor SVR lineal en el caso del experimento tipo 2 para cada combinación de LEDs 3. . . . .	77
A.33. Error medio absoluto para el regresor SVR lineal en el caso del experimento tipo 2 para cada combinación de LEDs. . . . .	77
A.34. Error relativo medio para el regresor SVR lineal en el caso del experimento tipo 1 para cada combinación de LEDs 3. . . . .	77
A.35. Error relativo medio para el regresor SVR lineal en el caso del experimento tipo 1 para cada combinación de LEDs 3. . . . .	77
A.36. Error relativo medio para el regresor SVR lineal en el caso del experimento tipo 2 para cada combinación de LEDs. . . . .	78
A.37. Error cuadrático medio para el regresor SVR polinomial en el caso del experimento tipo 1 para cada combinación de LEDs 1. . . . .	82
A.38. Error cuadrático medio para el regresor SVR polinomial en el caso del experimento tipo 1 para cada combinación de LEDs 2. . . . .	82
A.39. Error cuadrático medio para el regresor SVR polinomial en el caso del experimento tipo 2 para cada combinación de LEDs 1. . . . .	83
A.40. Error cuadrático medio para el regresor SVR polinomial en el caso del experimento tipo 2 para cada combinación de LEDs 2. . . . .	83
A.41. Error medio absoluto para el regresor SVR polinomial en el caso del experimento tipo 1 para cada combinación de LEDs 1. . . . .	83
A.42. Error medio absoluto para el regresor SVR polinomial en el caso del experimento tipo 1 para cada combinación de LEDs 2. . . . .	83
A.43. Error medio absoluto para el regresor SVR polinomial en el caso del experimento tipo 2 para cada combinación de LEDs 1. . . . .	83
A.44. Error medio absoluto para el regresor SVR polinomial en el caso del experimento tipo 2 para cada combinación de LEDs 2. . . . .	84
A.45. Error relativo medio para el regresor SVR polinomial en el caso del experimento tipo 1 para cada combinación de LEDs 1. . . . .	84

A.46.	Error relativo medio para el regresor SVR polinomial en el caso del experimento tipo 1 para cada combinación de LEDs 2. . . . .	84
A.47.	Error relativo medio para el regresor SVR polinomial en el caso del experimento tipo 2 para cada combinación de LEDs 1. . . . .	84
A.48.	Error relativo medio para el regresor SVR polinomial en el caso del experimento tipo 2 para cada combinación de LEDs 2. . . . .	84
A.49.	Error medio absoluto para el regresor SVR polinomial en el caso del experimento tipo 1 para cada combinación de LEDs 3. . . . .	85
A.50.	Error medio absoluto para el regresor SVR polinomial en el caso del experimento tipo 2 para cada combinación de LEDs 3. . . . .	85
A.51.	Error medio absoluto para el regresor SVR polinomial en el caso del experimento tipo 2 para cada combinación de LEDs. . . . .	85
A.52.	Error relativo medio para el regresor SVR polinomial en el caso del experimento tipo 1 para cada combinación de LEDs 3. . . . .	85
A.53.	Error relativo medio para el regresor SVR polinomial en el caso del experimento tipo 2 para cada combinación de LEDs 3. . . . .	85
A.54.	Error relativo medio para el regresor SVR polinomial en el caso del experimento tipo 2 para cada combinación de LEDs. . . . .	86
A.55.	Error cuadrático medio para el regresor SVR rbf en el caso del experimento tipo 1 para cada combinación de LEDs 1. . . . .	90
A.56.	Error cuadrático medio para el regresor SVR rbf en el caso del experimento tipo 1 para cada combinación de LEDs 2. . . . .	90
A.57.	Error cuadrático medio para el regresor SVR rbf en el caso del experimento tipo 2 para cada combinación de LEDs 1. . . . .	91
A.58.	Error cuadrático medio para el regresor SVR rbf en el caso del experimento tipo 2 para cada combinación de LEDs 2. . . . .	91
A.59.	Error medio absoluto para el regresor SVR rbf en el caso del experimento tipo 1 para cada combinación de LEDs 1. . . . .	91
A.60.	Error medio absoluto para el regresor SVR rbf en el caso del experimento tipo 1 para cada combinación de LEDs 2. . . . .	91
A.61.	Error medio absoluto para el regresor SVR rbf en el caso del experimento tipo 2 para cada combinación de LEDs 1. . . . .	91
A.62.	Error medio absoluto para el regresor SVR rbf en el caso del experimento tipo 2 para cada combinación de LEDs 2. . . . .	92

A.63.	Error relativo medio para el regresor SVR rbf en el caso del experimento tipo 1 para cada combinación de LEDs 1. . . . .	92
A.64.	Error relativo medio para el regresor SVR rbf en el caso del experimento tipo 1 para cada combinación de LEDs 2. . . . .	92
A.65.	Error relativo medio para el regresor SVR rbf en el caso del experimento tipo 2 para cada combinación de LEDs 1. . . . .	92
A.66.	Error relativo medio para el regresor SVR rbf en el caso del experimento tipo 2 para cada combinación de LEDs 2. . . . .	92
A.67.	Error medio absoluto para el regresor SVR rbf en el caso del experimento tipo 1 para cada combinación de LEDs 3. . . . .	93
A.68.	Error medio absoluto para el regresor SVR rbf en el caso del experimento tipo 2 para cada combinación de LEDs 3. . . . .	93
A.69.	Error medio absoluto para el regresor SVR rbf en el caso del experimento tipo 2 para cada combinación de LEDs. . . . .	93
A.70.	Error relativo medio para el regresor SVR rbf en el caso del experimento tipo 1 para cada combinación de LEDs 3. . . . .	93
A.71.	Error relativo medio para el regresor SVR rbf en el caso del experimento tipo 2 para cada combinación de LEDs 3. . . . .	93
A.72.	Error relativo medio para el regresor SVR rbf en el caso del experimento tipo 2 para cada combinación de LEDs. . . . .	94



# Acrónimos

**AOA** Angle-of-Arrival. 10

**BP** Back Propagation. 12, 15

**CDF** Cumulative Distribution Function. 41–43, 46, 48, 51, 55

**FDA** Fisher Discriminant Analysis. 12, 15

**GPS** Sistema de Posicionamiento Global. 1, II, V

**HOG** Histogram of Oriented Gradients. 18

**kernel RBF** Radial basis function kernel. 25

**KRRML** Kernel Ridge Regression ML. 8, 9, 15, XV

**LBP** Local Binary Patterns. 18

**LED** Light Emitting Diode. 2–4, 7–15, 28–36, 39–53, 55, 57–62, 64, II, III, V, VI, X–XII, XV, XVI

**LR** Linear Regression. 22, 23

**LRML** Linear Regression ML. 8, 9, 11, 15, XV

**LS** Least squares. 25, 36, 61

**L-SVM** Linear Support Vector Machine. 12, 15

**MAE** Error medio absoluto. 37, 44, 48, 53, 57, III, VI, VIII

**MAER** Error medio relativo. 38, 44, 48, 53, 57, III, VI, VIII

**ML** Machine Learning. 2, 3, 6, 7, 10, 11, 13, 15–17, 19, 22, 28, 60, 64, II, V, XV

**MLP** Multilayer Perceptron based regressor. 9, 12

**MRL** Multiple RL. 22

**ODS** Objetivos de Desarrollo Sostenible. 62

**QSC** Quadrant-Solar-Cell. 10

**RBF** Radial basis function. 4, 15, 35, 49, 53, 54, 57, 58, 61, 98, III, VI, VIII, XII, XIV

**RL** Reinforcement Learning. 19

**RMSE** Error cuadrático medio. 37, 41–44, 46, 51, 52, 55, 57–59, 61, III, VI, VIII

**RRML** third order ridge regressor. 10, 11, 15

**SFDP** Sigmoid Function Data Preprocessing. 8, 9, 15, XV

**SL** Supervised Learning. 17, 18, 22, 23

**SRL** Simple RL. 22

**SVR** Support Vector Regression. 4, 15, 23, 24, 35, 45, 48–50, 52–54, 57, 58, 61, 74, 78, 82, 86, 90, 94, 98, III, VI–IX, XI–XIV

**TFG** trabajo de Fin de Grado. 2–8, 14, 15, 25, 27–29, 33, 35, 39, 60–64, I, II, V, VI, X, XVI

**UL** Unsupervised Learning. 18

**UWB** Ultra Wide Band. 7

**VLP** Visible Light Positioning. 3, 7, 9–11, 13–15, VII, XV

**WK-NN** Weighted K-Nearest Neighbour. 9, 15

# Capítulo 1

## Introducción

En los últimos años, se han implantado distintas tecnologías como medida para la obtención de las coordenadas en una posición debido a todos los beneficios que se podrían conseguir. Uno de los sistemas más extendidos es el Sistema de Posicionamiento Global (GPS), pero su uso bajo ciertas condiciones es limitado. Esta tecnología, al igual que muchas otras utilizadas en interiores o ciertos entornos cerrados, no permite su correcto funcionamiento [1]. Ya que, por ejemplo, no se puede utilizar donde no se tenga acceso directo de las señales de los satélites. Así, a medida que la señal se propaga a través de medios, materiales y superficies su intensidad, se va atenuando pudiendo causar problemas de pérdida total de la intensidad de la señal. Además, las ondas electromagnéticas pueden sufrir efectos de reflexión y refracción al pasar de un medio a otro, dependiendo de sus respectivos índices de refracción, con la consecuentes pérdidas de potencia. De este modo, se produce más pérdida de la señal, causando que esta tecnología no sea fiable y precisa para su uso en interiores [2].

En base a esto, se han buscado diferentes sistemas de localización con los que poder estimar la ubicación o posicionamiento de un objeto de interés en interiores teniendo en cuenta las condiciones mencionadas de atenuación de la señal. Además, el despliegue se tiene que llevar a cabo en un entorno cerrado en el que el espacio es limitado, por lo que los dispositivos que formen parte del sistema deben tener un tamaño adecuado, ya que de lo contrario no se podrán instalar e implementar para su utilización, un ejemplo se ve en la Figura 1.1.

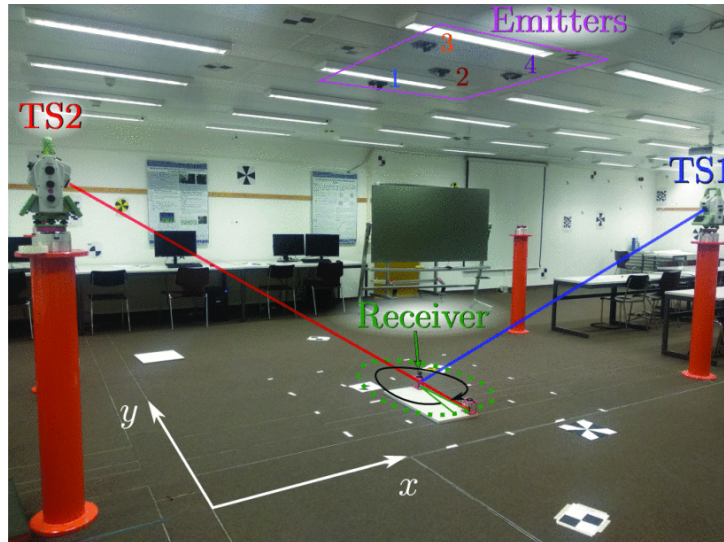


Figura 1.1: Escenario con la implantación de sensores y luces Light Emitting Diode (LED) que muestra el espacio necesario para un correcto comportamiento del sistema [3].

Uno de los sistemas que actualmente está experimentando un gran avance en los últimos años es el posicionamiento a través de la luminancia o usando ondas de luz. Por un lado, el desarrollo de luces de bajo consumo permite una mayor eficiencia energética, lo que conlleva un ahorro económico. Por tanto, un sistema de posicionamiento basado en ondas lumínicas permite una despliegue eficiente energéticamente y sin una gran inversión previa [4].

Finalmente, debido a las características necesarias de los modelos que se necesitaban para la estimación en interiores y las posibilidades que ofrecían las luces LEDs, se llegó a la conclusión de que se podrían utilizar este tipo de luces para llevar a cabo esta tarea. De tal manera que, con la instalación de luces LEDs de diferente frecuencia y un fotodiodo que recibiera las intensidades de cada LED se podría inferir la posición del objeto de interés analizando las diferencias de intensidad. En este Trabajo de Fin de Grado (TFG) se ha utilizado Machine Learning (ML) para tratar de realizar la estimación del posicionamiento en función de la luz recibida, debido a los buenos resultados que ha experimentado en los últimos años y lo necesario que es para este sistema [5]. Para ello, después del entrenamiento de los algoritmos de regresión con una base de datos, se podrían introducir los datos de los sensores de luminancia en estos regresores y, así obtener una estimación de la posición calculada.

En este TFG se ha utilizado una base de datos con distintas posiciones en un espacio y con las respectivas intensidades recibidas en cada posición por la luz de unos LEDs colocados de forma controlada, para estudiar cómo funcionan diferentes algoritmos de regresión y cómo de precisos son a la hora de realizar la estimación de las posiciones en el plano.

## 1.1. Motivación

A lo largo de los años, la tecnología ha ido desarrollándose en gran medida con grandes avances como la digitalización [6] o la implementación de redes de comunicación con el 5G [7]. Así, también destaca el ML, que en los últimos años ha permitido llevar a cabo diferentes procesos antes impensables [8]. Esto se suma a que cada vez más, las distintas tecnologías que existen se encuentran al alcance de cualquiera debido a su bajo precio y a la poca necesidad de inversión en nuevos dispositivos [9].

De este modo, se pueden ir aplicando esos desarrollos en cada vez más situaciones. Como es el caso de este TFG, en el que mediante luces LEDs, un fotodiodo y diferentes algoritmos de regresión se busca poder estimar la posición en interiores. Para ello se desea evaluar la eficiencia de esta tecnología en caso de que se quisiera aplicar su uso. Entre algunos ejemplos, la localización en interiores se podría aplicar en el sector industrial, ayudando a calcular la ubicación de robots autónomos que podrían moverse sin la necesidad de un operario que los controle [10]. También, al ser una tecnología que no necesita un gran desembolso económico, podría aplicarse a nivel usuario final en los hogares de las personas, ya sea para emplearlo en la ubicación para dispositivos como aspiradoras [11] o en personas con poca visión [12]. Esta tecnología permitiría saber dónde se encuentran algunos objetos o incluso un usuario dentro de un entorno acotado [13].

Como se puede ver, la aplicación de esta tecnología tiene varios usos. Algunos pueden ayudar a nivel laboral y económico a empresas, permitiendo tanto un ahorro de costes como mejorar la eficiencia de ciertos procesos. Además, permitiría a muchas personas mejorar su vida diaria en gran medida mediante dispositivos muy sencillos de instalar. Todas las aplicaciones nombradas se encuentran representadas en la Tabla 1.1.

Tabla 1.1: Listado de las aplicaciones nombradas en las que se puede aplicar la tecnología LED con el ML.

<b>Algunas aplicaciones de Visible Light Positioning (VLP) con ML</b>
Permitir el movimiento de robots autónomos en el sector logístico [10].
Facilitar a usuarios a ubicarse dentro de un edificio o a ubicar objetos [13].
Posibilitar a personas con una discapacidad visual a moverse en lugares cerrados [12].
Mejorar el nivel de vida en el hogar mediante la capacitación de dispositivos a moverse de forma autónoma [11].

## 1.2. Objetivos

El objetivo principal de este TFG es diseñar, implementar y evaluar una serie de algoritmos de ML para la estimación del posicionamiento de un objeto de interés a partir

de la luminancia recibida a través de múltiples fuentes de luz. A continuación, se exponen los objetivos parciales de este TFG:

- Se desarrollarán varios algoritmos de regresión, en particular, la regresión lineal y la regresión Support Vector Regression (SVR) con los kernel lineal, polinomial y Radial basis function (RBF) con el fin de estudiar su capacidad para predecir el posicionamiento.
- Se realizarán experimentos para determinar si es mejor predecir la posición estimando sus coordenadas cartesianas directamente, o por el contrario, las distancias radiales a las fuentes de luz y con estos últimos datos, mediante la intersección de circunferencias, obtener las coordenadas cartesianas.
- Usar varios LEDs para probar cómo afecta el número de LEDs usados en la estimación y, así determinar cuál es la cantidad de fuentes de luz necesarias para posicionar un objeto.
- Llevar a cabo los experimentos con distintos algoritmos de regresión, realizar un estudio comparativo de cada uno de los algoritmos para ver el rendimiento de cada uno y determinar cuál de los explorados es el que mejor estima las posiciones de los objetos de interés.

### 1.3. Estructura de la memoria

La memoria de este TFG se ha dividido en seis capítulos. De los cuales los Capítulos 1, 2 y 3 sirven para dar contexto del trabajo realizado. Por su parte, en los Capítulos 4, 5 y 6 se explican los experimentos llevados a cabo, así como los resultados y las conclusiones obtenidas. Además, al final se encuentran los apartados de Bibliografía y Anexo. A continuación, se presentan brevemente cada uno de los capítulos:

- Capítulo 1: Introducción. Se hace una descripción del tema tratado en este TFG. También se establecen los objetivos del mismo junto con la motivación de porqué se ha llevado a cabo.
- Capítulo 2: Estado del arte. Tras una investigación en literatura científica como artículos o libros, se expone la información recabada. Teniendo por un lado, la explicación de diferentes tecnologías relacionadas con el tema de este TFG y, por otro, distintas investigaciones y experimentos realizados por otras personas junto con sus resultados y conclusiones.

- Capítulo 3: Estudio teórico. En este capítulo, se exponen diferentes conceptos teóricos de las tecnologías que se van a utilizar más adelante. Además, se realiza una explicación teórica de distintas técnicas que se aplicarán a lo largo del TFG.
- Capítulo 4: Metodología. Aquí, se presenta la base de datos usada. También se explica todo el proceso realizado durante el desarrollo del TFG y cómo se han obtenido los distintos resultados.
- Capítulo 5: Resultados. En este capítulo, se incluyen todas las tablas y gráficas obtenidas de la realización de los distintos experimentos. Así como, un análisis de esos resultados.
- Capítulo 6: Conclusiones y Líneas Futuras. Finalmente, se desarrollan las conclusiones a las que se ha llegado la realización del TFG. También se presenta el impacto que puede tener este estudio y posibles líneas futuras de investigación.

# Capítulo 2

## Estado del arte

### 2.1. Introducción

Hoy en día, en un mundo en el que la tecnología avanza sin descanso, vemos cada vez más formas de ponerla a nuestro servicio. Así, entra en juego el uso del aprendizaje máquina, ML. Éste se puede utilizar para multitud de procesos mediante la aplicación de diferentes algoritmos, entre los que se encuentran la detección de rostros [14], comprensión de textos [15] y hasta para el desarrollo de coches autónomos [16]. En el contexto de este TFG, se aplicará ML para el cálculo del posicionamiento de objetos.

Antes de comenzar a desarrollar este trabajo, es necesario abordar el origen del ML [17] y cómo ha sido su evolución a lo largo de la historia, así se entenderá mejor la complejidad de este y algunos de sus usos. Uno de los primeros avances significativos en este campo fue realizado por Alan Turing [18], matemático que desarrolló el “test de Turing“ en la década de 1950. Este test consistía en determinar si una máquina era capaz de engañar a un humano haciéndolo creer que no era un máquina sino un humano y, por tanto, poseía inteligencia. Posteriormente, Arthur Samuel [19] creó un programa para jugar al ajedrez con una máquina, con la peculiaridad que ésta aprendía del jugador en cada partida. Otros grandes descubrimientos fueron la primera red neuronal inventada por Karl Steinbuch [20] o la red neuronal desarrollada por Frank Rosenblatt en conjunto con Charles Wightman para el reconocimiento de imágenes llamada Perceptrón [21].

En relación a la estimación de la localización o posicionamiento de objetos, existen diferentes tecnologías a tener en cuenta que a continuación serán explicadas. Una de ellas, es el uso de ultrasonido para obtener el posicionamiento y la orientación de un objeto [22]. Estos sistemas están compuestos por uno o varios generadores de ultrasonidos y uno o varios receptores de ultrasonidos. Para calcular la posición del objeto, se analiza tanto el tiempo de llegada de la señal emitida desde el generador al receptor como la diferencia entre el tiempo de llegada de distintas señales, si hay más de un generador o receptor.



Otra tecnología usada para la detección de la posición es Ultra Wide Band (UWB) [23]. De tal forma que con una serie de antenas dispuestas tanto en un emisor como en un receptor, se envían señales desde el emisor al receptor. Después de llegar al receptor se envían de vuelta al emisor y se calcula el tiempo que tardan las señales desde que el emisor las envía hasta que las recibe de vuelta y, así, determinar la distancia. Esto se usa, debido a que se sabe que las señales viajan a la velocidad de la luz. Además, UWB es muy utilizada para la estimación de la posición en interiores [24], así como las siguientes tecnologías.

Siguiendo el principio de la tecnología mencionada anteriormente se encuentra el WiFi [25]. Así, se calcula el tiempo que se tarda en recibir una señal en un dispositivo desde que se envió. Es importante mencionar que se necesitan varias antenas emisoras para poder establecer la posición correctamente.

Para obtener la posición de un objeto, también se utiliza el Bluetooth [26]. Este se basa en conectar dos dispositivos mediante Bluetooth y, analizando la intensidad de la señal entre los dispositivos, calcula la distancia exacta a la que se encuentran. En este caso, también se necesita más de un dispositivo con el que conectarse para obtener una posición concreta.

Como se ha mostrado, existen diferentes formas de obtener el posicionamiento de un objeto. En este TFG se usará una tecnología basada en VLP [27]. VLP se basa principalmente en una fuente de luz, como una bombilla LED, y un receptor que puede ser un fotodiodo [28] o un sensor de imagen formado por multitud de fotodiodos [29]. Normalmente, se utilizan varios LEDs, cada uno emitiendo luz a una frecuencia distinta. De esta forma, los receptores están expuestos a la luz y captan la intensidad procedente de cada LED. Tras los fotodiodos, hay una serie de filtros paso banda que separan la intensidad de cada LED por su frecuencia. De esta forma, se obtiene una base de datos con información correspondiente a cada LED. Después, estos datos pasan por un proceso de análisis gracias al ML y, mediante diferentes regresores y algoritmos se puede obtener una estimación de la posición exacta de cada objeto.

## 2.2. Trabajos de VLP

Para entender la situación actual de VLP se ha realizado una búsqueda de diferentes trabajos de investigación que emplean este sistema. Para ello, se ha usado la plataforma Google Scholar y se han elegido 7 artículos para su análisis detallado. Para lo cual, se ha usado como *keyword* «*visible light positioning*». Para el filtrado y selección de los diferentes artículos se ha seguido el criterio de haberse publicado en el año 2018 o posterior, tener alguna característica definitoria que le haga parecerse a este trabajo y haber sido citado

varias veces demostrando, de esta forma, su relevancia. A continuación, se procederá a explicar los trabajos seleccionados.

Una de estas investigaciones aparece en el artículo [30], en él se lleva a cabo un proceso muy parecido al que más tarde se desarrollará en este TFG. En este estudio, tenemos tres LEDs colocados de forma triangular sobre un espacio cada uno emitiendo luz a frecuencias distintas de 23kHz, 53kHz y 83kHz. Estos LEDs tienen las coordenadas  $x, y, z$  siguientes: (0cm, 0cm, 0cm), (25cm, 43.3cm, 0cm) y (50cm, 0cm, 0cm). A continuación, se colocan diferentes objetos, cada uno con un fotorreceptor que respectivamente recibirá la luz de cada LED. Además, para el experimento se plantean cuatro distancias o alturas diferentes en las que poner los objetos con los fotorreceptores. Estas distancias son diferentes para los grupos de entrenamiento y para los de test, siendo las del grupo de entrenamiento 80cm, 100cm, 120cm y 150cm; y las de test 95cm, 100cm, 115cm y 135cm. Por otro lado, se proponen 16 posiciones como posibles ubicaciones para el grupo de entrenamiento y 30 para el de test, midiéndose la intensidad en cada posición 20 veces. Asimismo, el área permitida para la posición de los objetos es de 50cm x 50cm. Tras esto, se procede al análisis de las intensidades según las distintas frecuencias para obtener cada una separada. Esto se hace separando cada frecuencia con distintos filtros paso banda. De este modo, obtenemos una serie de intensidades clasificadas por su frecuencia que podemos relacionar con unas coordenadas dentro del espacio. En este experimento, se quiso comparar también el uso de diferentes algoritmos junto con un método de preprocesamiento y ver si había algún cambio. Los algoritmos usados fueron el algoritmo de Linear Regression ML (LRML) y Kernel Ridge Regression ML (KRRML). Además, cada uno de ellos se aplica a dos grupos: uno con los datos obtenidos sin modificar y otro aplicando Sigmoid Function Data Preprocessing (SFDP).

Tras el experimento se dedujo que con el preprocesado de los datos se obtenían resultados más precisos, así como que con KRRML se reducía el error en comparación con LRML. En la Tabla 2.1, se puede observar el error correspondiente a cada caso. Por último, destacar que este conjunto de tres LEDs se puede replicar a lo largo de un espacio mayor siempre y cuando los LEDs que usan la misma frecuencia estén lo suficientemente alejados para que no interfieran unos con otros.

Method	Ave. horizontal error	Ave. vertical error
LRML	3,74 cm	3,64 cm
LRML with SFDP	2,70 cm	2,84 cm
KRRML	2,00 cm	2,23 cm
KRRML with SFDP	1,96 cm	2,16 cm

Tabla 2.1: Resumen de los errores medios obtenidos en [30] para los algoritmos LRML, KRRML con y sin SFDP.

Siguiendo con otras investigaciones se encuentra otro artículo [29], en el que se aborda este problema desde otra perspectiva. Así, en cada posición se utilizará más de un fotodiodo. El diseño de este pasaría a un receptor con forma triangular y en cada lado iría un fotodiodo. Estos lados estarían inclinados  $60^\circ$  y, además, en el centro en horizontal encontraríamos un cuarto fotodiodo representado en la Figura 2.1.

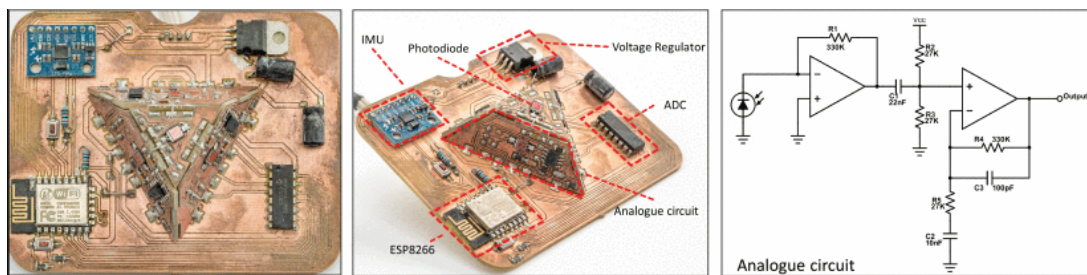


Figura 2.1: Diseño del receptor utilizando en [29] donde se agrupan cuatro fotodiodos.

Esta forma de medición evitaría tanto los problemas de inclinación como la pérdida de visión directa del LED. En cuanto al número de LEDs, en este caso, el experimento cuenta con cuatro LEDs, cada uno emitiendo a una frecuencia distinta siendo estas de valor 2,5kHz, 2,7kHz, 3,2kHz y 3,5kHz. La distancia entre los LEDs y plano donde se encuentran los objetos a posicionar hay 160cm. Además, se escogen 169 posiciones para un espacio de 1200cm x 1200cm. En cada posición se toman 32 muestras y como grupo de entrenamiento se toman 32 posiciones, cada una con hasta 16 intensidades distintas correspondientes a los cuatros LEDs y a los cuatro fotodiodos, quedando el resto para el grupo de prueba. Por otro lado, como algoritmo usa Weighted K-Nearest Neighbour (WK-NN).

Este muestra un buen desempeño que, comparado con Multilayer Perceptron based regressor (MLP), sigue siendo mejor que este último. Además, para este experimento se hacen diferentes pruebas comparando VLP con una solución comercial dedicada a la localización como es *HTC Vive*. *HTC Vive* es una tecnología de captura de movimiento usada

principalmente para realidad virtual. Esta solución consiste en dos focos que transmiten haces de luz y un dispositivo que sirve para registrar esa luz y determinar la posición, tal y como se aprecia en la Figura 2.2. Los resultados obtenidos entre ambos métodos son similares, siendo el error medio del valor de 4,74mm.

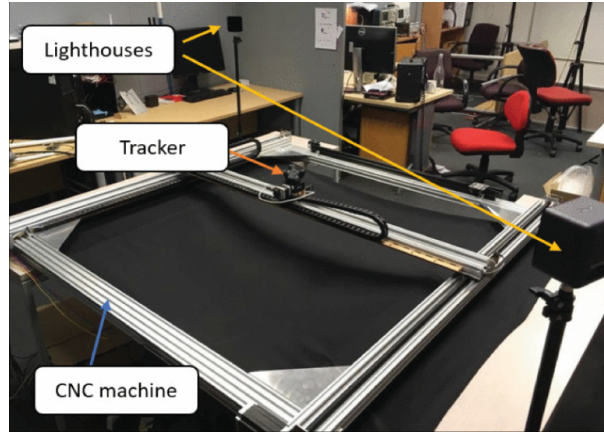


Figura 2.2: Montaje tecnología HTC Vive utilizado en [29].

Continuando con otros ejemplos, se encuentra la investigación [31]. Este trabajo también plantea el uso de VLP conjuntamente con un cuadrante de celdas solares Quadrant-Solar-Cell (QSC), véase la Figura 2.3. Este QSC, hecho de silicio, se encuentra dividido por dos barreras verticales de cierta altura y, tanto esta como su posición deben ser definidas desde un principio y no modificarse, ya que esta investigación se basa en la sombra que se produce por las barreras y la luz que llega a las celdas.

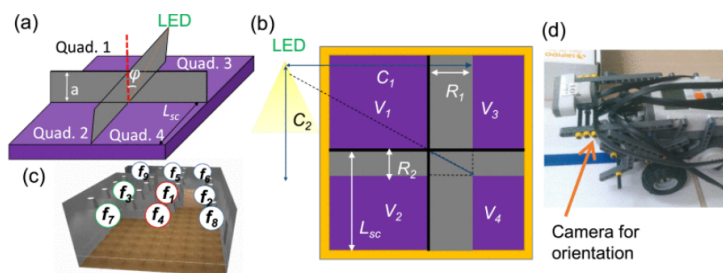


Figura 2.3: Diseño del cuadrante de celdas solares usado en [31].

Así, se usa el Angle-of-Arrival (AOA) como forma de medición haciendo una conversión óptico-eléctrica. De tal forma que, como grupo de entrenamiento tenemos un total de 25 posiciones en las que en cada una se obtiene el voltaje de cada cuadrante, las medidas de cada cuadrante y las medidas de la sombra causada, dando lugar a sets con 20 datos distintos. Como fuente de luz, se usa un LED y el tamaño de la celda es de 40cm x 40cm. Para probar la eficacia de este método, se comparan los resultados sin ML, con el regresor de orden 3 (*third order regressor*) y con Third order ridge regressor (RRML). Obteniéndose finalmente unos resultados significativamente más precisos aplicando ML,

y, a su vez siendo algo mejores con RRML que con el regresor ML de orden 3. Siendo los errores de 7,22cm sin ML, 3,20cm para *third order regressor* y 3,08cm para RRML.

Por otro lado, tenemos otro artículo con su correspondiente investigación [32]. En este caso, se prueba VLP junto con un software llamado *DIALux*. Su funcionamiento se puede observar en la Figura 2.4. Este se encarga de proporcionar información de cómo es el espectro que emite una bombilla. Así, tenemos un escenario muy similar al caso de nuestro proyecto. Primero, situados en el techo tenemos 4 LEDs que emiten luz con una frecuencia distinta de 47kHz, 59kHz, 83kHz y 101kHz. También, se usa un robot autónomo que es el que se sitúa en las diferentes posiciones y el que dispone en su parte superior de un fotodetector. Estas posiciones elegidas se han medido 30 veces cada una. Para el experimento, se usó una habitación de 200cm x 155cm y la altura desde el LED hasta el fotorreceptor es de 270cm. A continuación, se obtienen las diferentes medidas y se presentan dos casos: uno en el que las intensidades provienen del software *DIALux* y otro en el que provienen de las mediciones captadas con el fotodetector. De cada uno de estos grupos, se seleccionan unos valores para el grupo de entrenamiento y se les aplica LRML de forma independiente. Obteniendo dos estimaciones distintas.

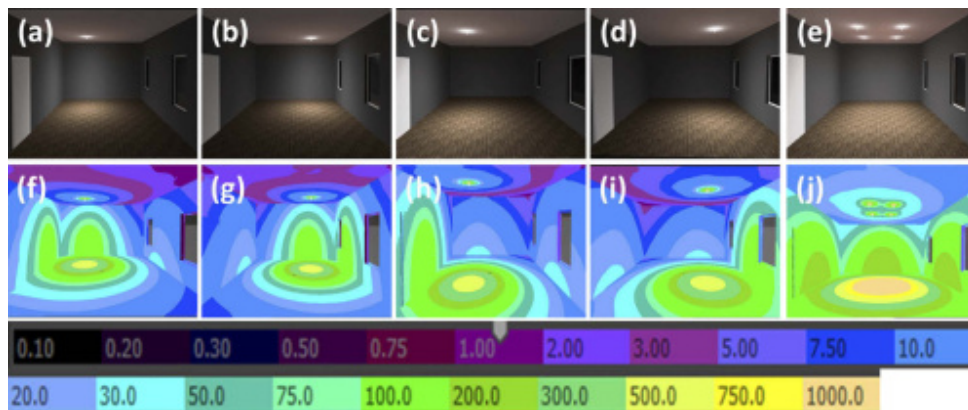


Figura 2.4: Funcionamiento del software *DIALux* [32].

Finalmente, se muestra que para ambos casos los resultados son muy similares teniendo una mayor dificultad para saber la posición cuanto más alejado esté el punto de los LEDs. Obteniendo un error de 10,5cm con los datos experimentales y 11,1cm con la simulación de *DIALux*.

Otra investigación a destacar es [33]. Para esta investigación, la diferencia que encontramos es que el receptor es una cámara que analiza el tamaño y la forma de la imagen que se forma con la luz que le llega del LED, que tiene una forma circular o de elipsoide según la posición de estos, como se observa en la Figura 2.5.

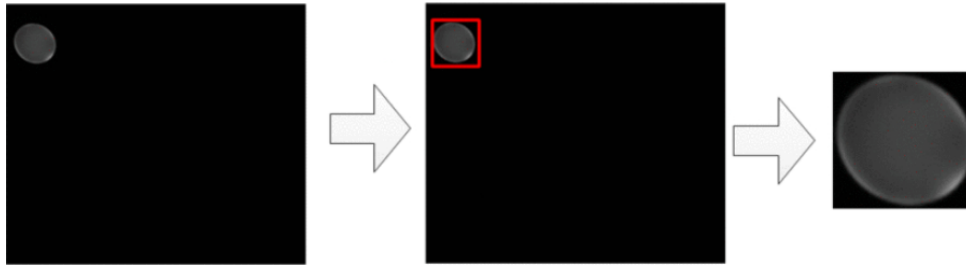


Figura 2.5: Segmentación de la imagen realizada en [33].

Los datos de entrada para el experimento son las coordenadas  $x$  e  $y$  del centro de la imagen, el área de esta y sus ejes mayor y menor. Como fuentes de luz, hay 5 LEDs y la distancia entre la cámara y los LEDs varía entre 10cm y 160cm, tomando 6 muestras cada 10cm para el grupo de entrenamiento. Como área para las posiciones se establecen las medidas de 130cm x 130cm. Además, para el análisis de los datos se usará MLP, en concreto, el algoritmo para el test que se utilizará es (Back Propagation (BP)) comparando los resultados tras el uso de redes neuronales con distintos valor de capas o nodos: 10, 30 y 50. Con todo esto, se consigue obtener con cierta precisión la posición en tiempo real, siendo el error medio 3,6cm, 2,1cm y 1,9cm, es decir, a mayor número de nodos mejor es el resultado.

Otra investigación analizada en este trabajo es [34]. En este caso, se busca identificar distintos LEDs mediante la luz que emiten y que una cámara recibe para analizar sus características, así después obtener la posición del receptor, como se observa en la Figura 2.6. En total, en el experimento se usan 15 LEDs y la superficie disponible tiene las medidas de 170cm x 80cm. En cada posición, se toman 10000 muestras, variando la distancia entre los LEDs y la cámara de 100cm hasta los 600cm. Para ello, cada LED emite luz a una frecuencia diferente y con diferente ciclo de trabajo. Esto queda representado en la cámara a través de una serie de imágenes circulares de distinto tamaño y con líneas blancas dependiendo de la frecuencia del LED. De tal forma que, teniendo en cuenta el número de rayas blancas, el área total de la imagen y el ciclo de trabajo de cada imagen y, aplicando diferentes algoritmos se estima la posición. Los algoritmos utilizados son el Fisher Discriminant Analysis (FDA) y Linear Support Vector Machine (L-SVM). Tras realizar el análisis, se observan con ambos métodos unos resultados bastante similares, variando el error entre 0,04cm y 0,007cm dependiendo de la distancia y del número de líneas blancas que tiene cada círculo.

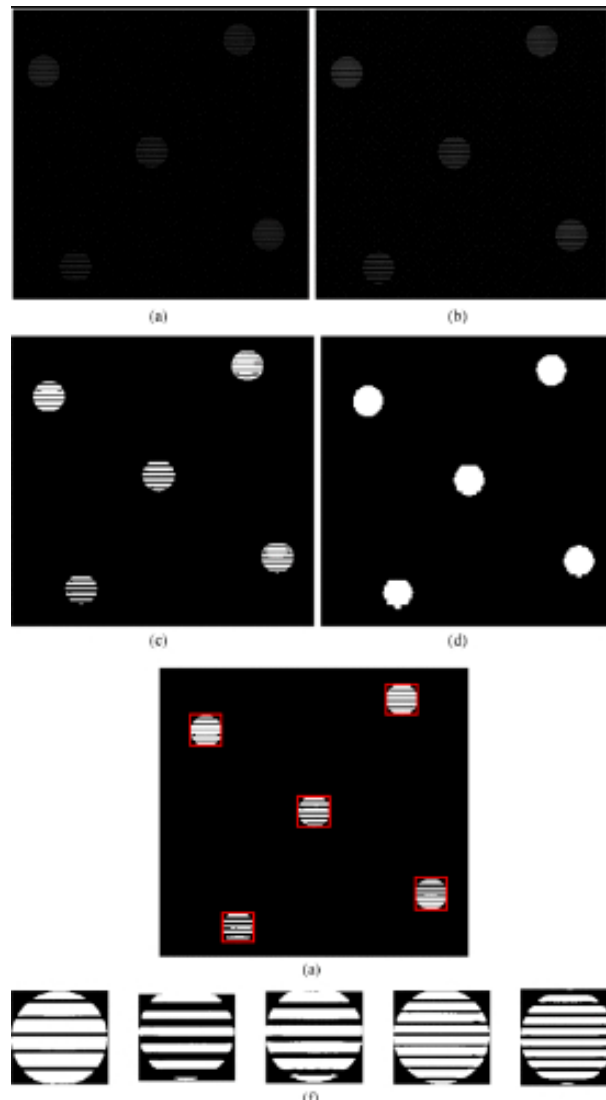


Figura 2.6: Segmentación de las imágenes obtenidas por la cámara [34].

Para finalizar con las diferentes investigaciones, cabe nombrar el estudio realizado en [35]. Su principal característica se basa en ver cómo afecta el uso de la interpolación lineal a la precisión de un modelo. En el experimento utilizaron 3 LEDs con frecuencias de 20kHz, 40kHz y 60kHz. Además, se realizaron diferentes experimentos con diferente número de muestras, en total tres grupos: un grupo reducido de 19 muestras de entrenamiento en las que cada posición se medía 30 veces, por lo que, finalmente se tenían 570 muestras. Otro grupo con un mayor número de muestras de entrenamiento, 45 muestras, obtenidas de aplicar el algoritmo de interpolación lineal al anterior grupo. Por último, un tercer grupo con el mayor número de muestras de los tres mencionados, 81 muestras, a las cuales se les aplica el modelo de regresión de segundo orden de ML (*second order regression ML model*), obteniendo un total de 2430 muestras. El área usada para las posiciones es de 50cm x 45cm. Tras el experimento, se puede deducir que, tanto para el grupo con el mayor número de muestras como el grupo obtenido con la interpolación se obtienen resultados muy similares. Quedando el grupo de menor muestras opacado totalmente por estos otros

dos grupos en cuanto a precisión a la hora de obtener la posición. Siendo los siguientes resultados los que representan el funcionamiento de este experimento: en el caso de 19 muestras, el 70 % de las muestras han obtenido un error menor de 4cm. Para el caso de 45 muestras, un 83 % de las mismas ha obtenido un valor menor a 4cm de error. Por último, para las 81 muestras, el porcentaje de muestras que ha obtenido un error menor de 4cm es 87 %.

Tras el estudio y análisis de las mencionadas investigaciones, se ha realizado un resumen de las características más importantes de cada investigación. Esta información se ha organizado en las Tablas 2.2 y 2.3.

Tabla 2.2: Resumen de los trabajos analizados sobre VLP 1.

Investigación	Nº LEDs	Nº Muestras	Distancia entre LED y posición	Superficie prueba
[30]	3	Entrenamiento: 16 Test: 30	Entrenamiento: 80, 100, 120 y 150 cm Test: 95, 100, 115 y 135 cm	50cm x 50cm
[29]	4	Entrenamiento: 32 Test: 137	160 cm	1200cm x 1200cm
[31]	1	Entrenamiento: 25 con 20 datos cada una	—	40cm x 40cm
[32]	4	—	270cm	200cm x 155cm
[33]	5	Entrenamiento: 36	10 cm -160cm	130cm x 130cm
[34]	15	Entrenamiento : 10000	100cm - 600cm	170cm x 80cm
[35]	3	Entrenamiento: 570, 45 y 2430	—	50cm x 45cm
TFG	4	Entrenamiento: 126 Test: 32	128,4cm	300cm x 300cm



Tabla 2.3: Resumen de los trabajos analizados sobre VLP 2.

Investigación	Características propias de la investigación	Algoritmos	Errores
[30]	Ver el desempeño del sistema con el uso de un algoritmo de preprocesamiento	LRML KRRML LRML con SFDP KRRML con SFDP	LMRL: Vertical: 3,64 cm y Horizontal: 3,74cm LRML con SFDP: Vertical: 2,84cm y Horizontal: 2,10cm KRRML: Vertical: 2,23cm y Horizontal: 2,00cm KRRML con SFDP: Vertical: 2,16cm y Horizontal: 1,96cm
[29]	Uso de un receptor piramidal y comparar su desempeño con la tecnología HTC Vive	WK-NN	WK-NN: 4,74 cm HTC Vive: 4,74 cm
[31]	Aplicación de un cuadrante de celdas solares para la obtención de datos	3 <sup>o</sup> order Regressor RRML	Sin ML: 7,22cm 3 <sup>o</sup> order Regressor: 3,20cm RRML: 3,08cm
[32]	Probar el funcionamiento del software DIALux para obtener la posición de objetos	LRML	LRML: 10,5cm DIALux: 11,1cm
[33]	Análisis de las características de las imágenes que generan los LEDs con una cámara	BP: N=10, 30 y 50	BP - N=10: 3,6cm BP - N=30: 2,1cm BP - N=50: 1,9cm
[34]	Análisis de las distintas imágenes que la cámara recibe	FDA y L-SVM	0,7cm - 4,00cm
[35]	Ver como afecta el uso de la interpolación lineal a la precisión de un modelo	Linear Interpolation 2 <sup>o</sup> order Regressor ML Model	570: 70 % 4cm 45: 83 % 4cm 2430: 87 % 4cm
TFG	Análisis de cada uno de los algoritmos de regresión con dos tipos de experimentos Analizar los resultados según el número de LEDs	Regresor lineal Regresor SVR lineal Regresor SVR polinomial Regresor SVR RBF	Regresor lineal: 10,8cm Regresor SVR lineal: 33,5cm Regresor SVR polinomial: 10cm Regresor SVR RBF: 7,5cm

# Capítulo 3

## Estudio teórico

### 3.1. Machine Learning

El ML aparece como una rama dentro del campo de la inteligencia artificial [36]. Ésta se basa en que las máquinas puedan aprender y tomar decisiones basadas en datos. Se ha convertido en una herramienta fundamental para el desarrollo de ciertas aplicaciones. Su importancia radica principalmente en su capacidad para abordar problemas complejos. Actualmente, con el creciente volumen de datos disponibles y la capacidad de procesamiento cada vez mayor, el ML se ha convertido en una herramienta de gran valor para extraer información útil a partir de grandes conjuntos de datos [37].

El ML permite a las computadoras aprender de los datos de manera iterativa. Al alimentar un modelo con datos, el modelo analiza y encuentra patrones o relaciones. Estos patrones aprendidos se utilizan posteriormente para hacer predicciones o tomar decisiones sobre nuevos datos. A medida que se proporcionan más datos al modelo, éste puede mejorar tanto sus predicciones como su rendimiento. Esto se debe a que el modelo ajusta y optimiza sus parámetros internos a medida que se le presenta más información, lo que lleva a una mayor precisión en la toma de decisiones. En consecuencia, el ML permite que las máquinas aprendan de forma autónoma, se adapten a nuevos escenarios y mejoren su rendimiento a medida que encuentran más datos y realizan más iteraciones de entrenamiento [38].

Cabe mencionar la importancia de analizar si el ML es la herramienta adecuada para resolver el problema que se pretende abordar. Pues dada la propia naturaleza del mismo tiene ciertas limitaciones, las cuales son principalmente dos. En primer lugar, en problemas que requieran una alta precisión, el ML tiene un margen de error que variará según la cantidad de datos y la disparidad de los mismos. En segundo lugar, debido al procesamiento necesario para llevar a cabo la lectura de los datos, para generar los distintos modelos y la aplicación de estos modelos a los datos de test conlleva un tiempo que puede exceder

en algunos casos el tiempo límite necesario para el problema que se quiere solventar.

Antes de ver los diferentes paradigmas de ML es conveniente explicar el proceso llevado a cabo para la creación y desarrollo de un algoritmo de ML.

1. Hay que hacer una toma de datos y seleccionar de aquellos que pueden resultar relevantes para el algoritmo que se quiere desarrollar, ya que algunos pueden no ser necesarios en nuestro problema.
2. El siguiente paso es el preprocesado de los datos. Aquí, se preparan los datos para que el algoritmo pueda usarlos. Algunos cambios necesarios pueden ser ajustar el formato de los datos o eliminar cualquier dato erróneo. Además, otros ajustes pueden consistir en normalizar los datos o discretizarlos.
3. A continuación, se procede al entrenamiento del algoritmo, para lo cual se seleccionan una serie de datos de todo el conjunto, siendo estos los que forman el grupo de entrenamiento. Como su nombre indica, este conjunto se usa para alimentar el algoritmo y que este genere las relaciones necesarias para la consiguiente aplicación con otros datos.
4. Después, se testea el algoritmo. Con las relaciones e información que ha obtenido el algoritmo del grupo de entrenamiento, habría que probar que su funcionamiento es el correcto y, para ello, se usaría el segundo conjunto, el grupo de test, probando, así, el funcionamiento del mismo.
5. Finalmente, se ejecuta el algoritmo con los datos de los que se quiere obtener una predicción o una estimación [39].

Como se ha mencionado, la forma de aplicar ML es utilizando algoritmos, éstos se pueden clasificar en distintos tipos según si los datos que se usan como entrada cumplan ciertas características o no. Por un lado, se encuentra el aprendizaje supervisado (Supervised Learning (SL)), que se podría definir como una rama del ML donde se entrena un modelo utilizando datos de entrada etiquetados. Es decir, los datos que se tienen para que el algoritmo se ejecute tienen asociada un etiqueta, de tal forma que, el algoritmo no solo aprende y crea relaciones entre valores numéricos sino también entre las diferentes características de las que se tienen datos [40]. Este tipo de algoritmo se usa en dos casuísticas de problemas muy concretas, que son clasificación y regresión. En el caso de la clasificación presentada en la Figura 3.1, dados los datos de entrada etiquetados se busca que el algoritmo elija una de las posibles categorías u opciones que se tienen como posibles soluciones, para ello se basa en los patrones que ha generado durante el entrenamiento del algoritmo. Por un lado, tenemos distintos conjuntos de datos accesibles de manera global para la realización de análisis (UCM, SIRI-WHU y SAT-6) representados de diferentes

formas (Histogram of Oriented Gradients (HOG) y Local Binary Patterns (LBP)) siendo cada color una categoría en la que se ha clasificado cada dato. Así, se puede observar un ejemplo e clasificación [39]. Por esta razón, es necesario que el conjunto de datos de entrada sea lo suficientemente amplio y variado para que el algoritmo pueda establecer las relaciones necesarias entre los datos. Por el lado de la regresión, se tiene otro tipo de problema en el que, con unos datos también etiquetados, se busca un resultado numérico, es decir, se establecen relaciones entre variables independientes (los datos de entrada) y variables dependientes (los datos de salida).

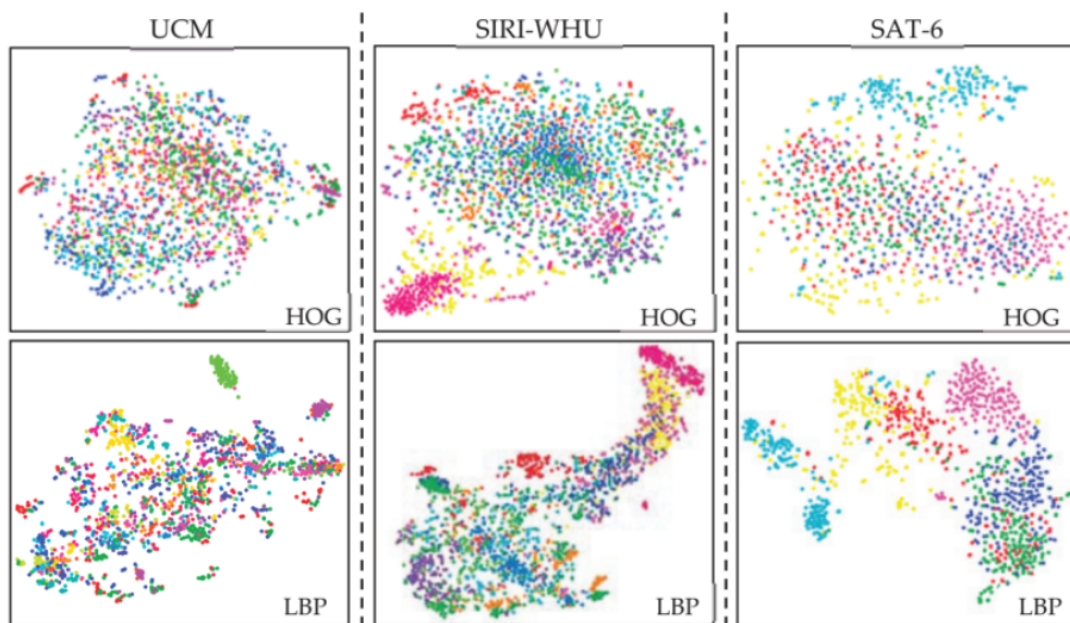


Figura 3.1: Ejemplo de algoritmo de clasificación de un problema de aprendizaje supervisado [39].

Por otro lado, la contraparte a SL es el aprendizaje no supervisado (Unsupervised Learning (UL)). Este paradigma de aprendizaje usa datos de entrada no etiquetados para establecer relaciones y segmentar los conjuntos de datos en función de similitudes o características comunes que a priori están ocultas. En este caso, no se tienen unas soluciones o salidas predefinidas a las que llegar. Un ejemplo de UL es el *clustering* presentado en la Figura 3.2 . Este problema se basa en establecer grupos entre los conjuntos de datos dadas ciertas características comunes entre los valores de los datos y en las relaciones entre estos datos [41].



Figura 3.2: Dos ejemplos de agrupación de datos en clusters [41].

Además de estos paradigmas de aprendizaje, existe otro a destacar, el aprendizaje por refuerzo (Reinforcement Learning (RL)). El RL se basa en la exploración mediante un serie de acciones y comportamientos realizados por un agente inteligente, cómo van evolucionando los resultados que se están obteniendo, a través de un sistema de recompensas. Este algoritmo se entrena a si mismo a través del ensayo y error, analizando las secuencias que se han seguido junto con sus resultados y otorgando recompensas a los secuencias correctas y penalizaciones a las incorrectas. Mientras el algoritmo se ejecuta, puede cambiar su política de que secuencias son más correctas y, así elabora una base de datos con las posibles acciones a seguir. Por esto último, es necesario establecer un balance entre el entrenamiento y número de iteraciones que puede hacer el algoritmo y el error que se puede permitir cometer [42].

## 3.2. Validación cruzada

Los algoritmos de ML tienen un margen de error a la hora de obtener las predicciones de los datos de prueba. Por ello, hay que realizar un estudio analizando estos errores, que reflejará la diferencia entre los datos estimados y los datos reales y, así, comprobar la eficacia de los distintos algoritmos [43].

Entre los distintos métodos que existen para comprobar la eficacia de los sistemas se encuentra la validación cruzada (*cross validation*). Un diagrama de la misma se presenta en la Figura 3.3. Esta es una técnica utilizada en el aprendizaje automático que consiste en evaluar y probar el rendimiento de un modelo, además de garantizar que los resultados son independientes de la elección que se realice entre los datos de entrenamiento y prueba. Para ello, en vez de tener un único conjunto de entrenamiento y de prueba, se realizan diferentes particiones de los datos formando cada una un subconjunto, así se utilizan diferentes combinaciones para ver el rendimiento del modelo. Finalmente, una vez se tienen todos los datos para cada subconjunto, se hace la media con todos los subconjuntos.

Existen diferentes técnicas de validación cruzada que pueden realizarse para la evaluación del modelo.

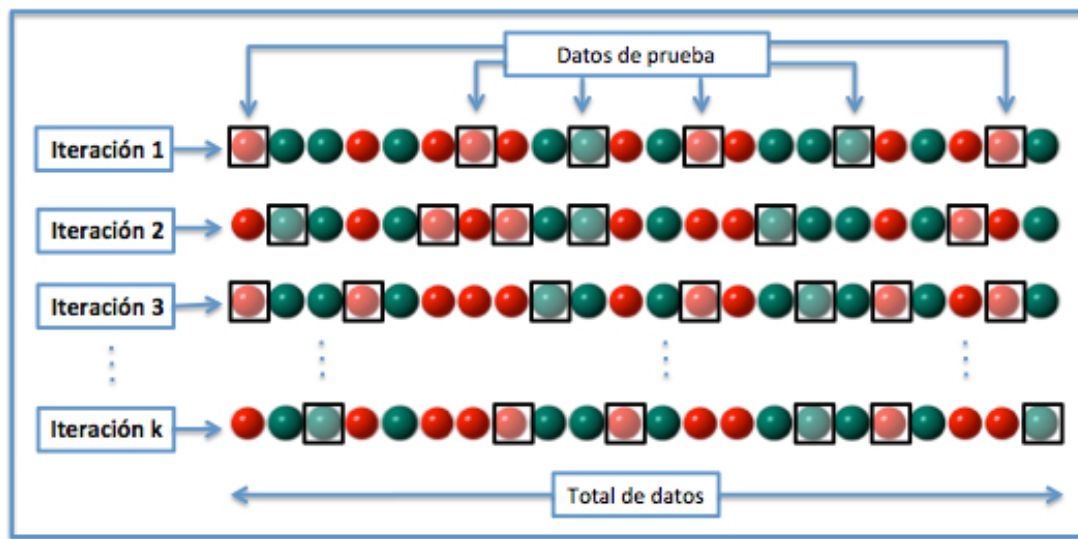


Figura 3.3: Diagrama resumen de la técnica de *Cross-validation* [44].

Una de ellas es *Single hold-out random subsampling*, presentada en la Figura 3.4. Es una de las técnicas más simples, ya que consiste en dividir el conjunto de los datos en dos grupos, uno para el grupo de entrenamiento y el resto para el de prueba. De los dos grupos, la mayoría de los datos, entre el 90 % y el 70 %, se usan en el grupo de entrenamiento y entre el 10 % y el 30 % para el grupo de prueba. Esta selección de los datos para cada conjunto se realiza de manera aleatoria.

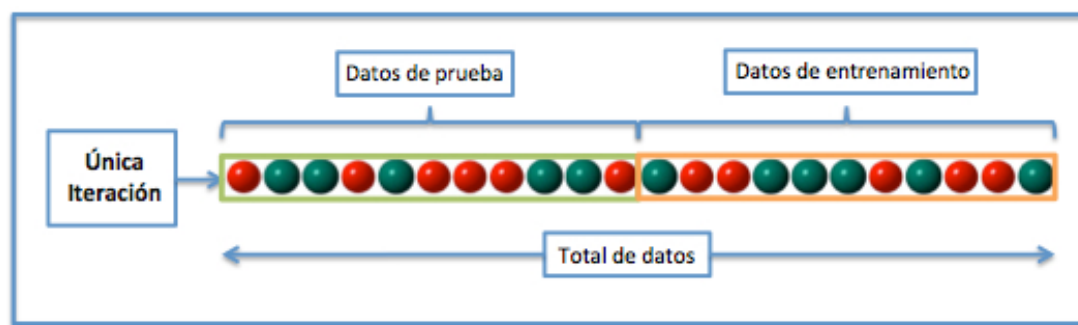


Figura 3.4: Diagrama explicativo de la técnica de *Single hold-out random subsampling* [44].

Otra de las técnicas más utilizadas es *k-fold*, presentada en la Figura 3.5, el proceso que se lleva a cabo en este caso es el de dividir el conjunto de datos en  $k$  subconjuntos y realizar  $k$  validaciones siendo el grupo de prueba para cada iteración un subconjunto distinto. El resto de subconjuntos en esa iteración pertenecerán al grupo de entrenamiento. Esta selección de los datos puede ser tanto de forma aleatoria como dividiendo todo el conjunto de datos de tal forma que cada dato unitario forme un subconjunto en si mismo.

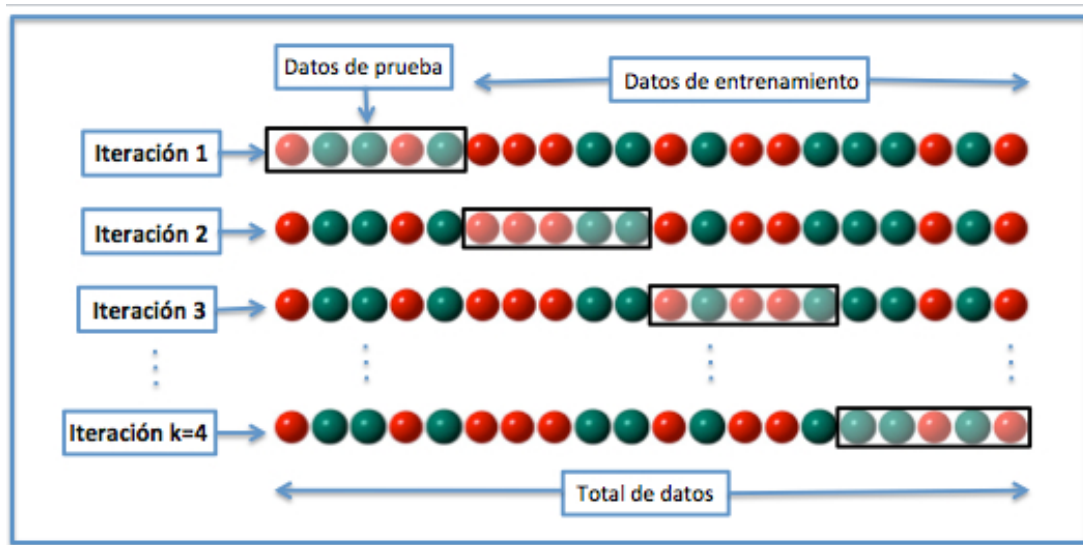


Figura 3.5: Diagrama explicativo de la técnica de *k-fold cross-validation* [44].

Dentro de la técnica *k-fold* existen diferentes variaciones. Una de ellas es la validación cruzada *leave-one-out*, vea la Figura 3.6, que produce tantos subconjuntos como número de datos, ya que cada subconjunto está formado por un solo dato. Este caso puede ser computacionalmente costoso, pero proporciona una evaluación muy precisa [45].

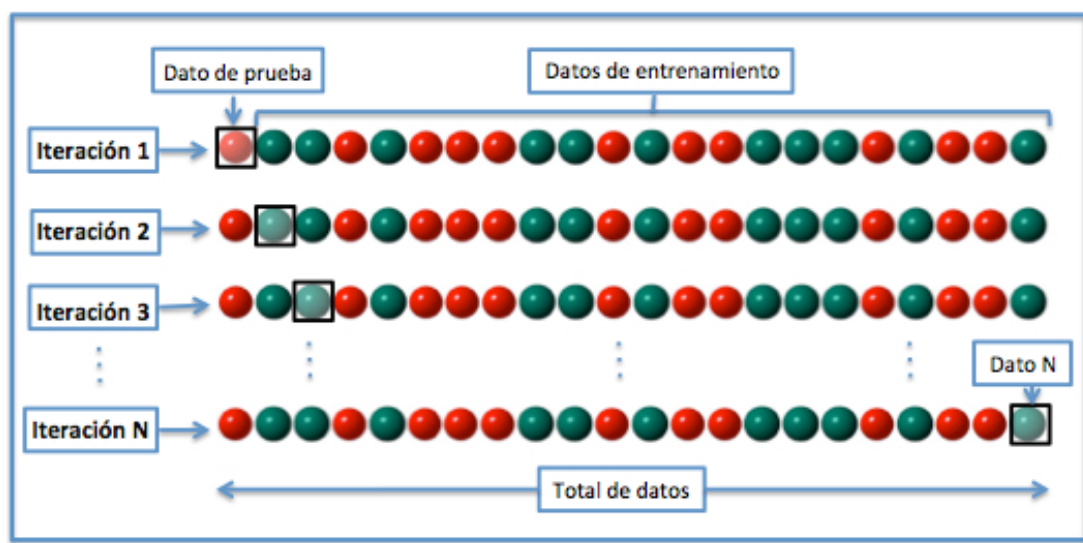


Figura 3.6: Diagrama explicativo de la técnica de *leave-one-out cross-validation* [44].

Otra técnica diferente es la validación cruzada estratificada (*Stratified Cross-Validation*). Esta es utilizada principalmente cuando los datos no están balanceados o están desequilibrados, es decir, tenemos distintas clases de datos, pero más cantidad de una clase que de otra. Por lo que los resultados podrían no ser correctos, ya que *k-fold* necesita que los datos estén divididos con una distribución de probabilidad uniforme.

Por último, existe otra técnica conocida como *Jackknife*. Esta técnica comparte algunas

características con la validación cruzada *leave-one-out*. En ambas técnicas, cada subconjunto está formado por un solo dato, pero en este caso al terminar todas las iteraciones se hace la media de todos los subconjuntos, para, después, calcular una estimación del *bias*, con la Ecuación (3.1) y de varianza, con la Ecuación (3.2).

$$bias(\hat{\theta}) = (n - 1)(\bar{\hat{\theta}} - \hat{\theta}) \quad (3.1)$$

$$Var(\hat{\theta}) = \frac{n - 1}{n} \sum_{i=1}^n (\hat{\theta}_i - \bar{\hat{\theta}})^2 \quad (3.2)$$

### 3.3. Regresores

Para la ejecución de esta investigación, en conjunto con las técnicas anteriores se han usado métodos de ML, entre ellos la regresión que entra dentro del SL. Así, se ha llevado a cabo un estudio del rendimiento de distintos algoritmos de regresión. Estos regresores se basan en un conjunto de funciones que dados unos datos de entrada permite estimar unos resultados. A continuación, se expondrán cada uno de estos regresores con sus características.

#### 3.3.1. Regresión Lineal

La Linear Regression (LR) es un modelo de regresión basado en una o varias entradas  $x_i$  independientes que, al aplicarles una función, dan lugar a una respuesta  $y_i$  siendo esta respuesta una línea recta y dependiente [46]. Este tipo de regresión no tiene gran complejidad y su ejecución es de las más simples, así como su tiempo de procesamiento es menor en comparación con otros regresores. Por esto mismo, sus resultados suelen ser menos precisos. Según el número de entradas que se tengan, tendremos regresión lineal simple Simple RL (SRL), Ecuación (3.3) o regresión lineal múltiple Multiple RL (MRL), Ecuación (3.4). Para una entrada, tendremos SRL y para dos o más entradas será MRL [47] [48] [49]. Asimismo, las ecuaciones para llevar a cabo la regresión variarán según el caso, haciendo uso de más variables y datos en el caso de MRL.

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i, i = 1, 2, \dots, n \quad (3.3)$$

$$y_{i,j} = \beta_0 + \sum \beta_j x_{i,j} + \epsilon_i, i = 1, 2, \dots, n \quad (3.4)$$

En la que  $y_i$  es el valor calculado, es decir, la respuesta estimada,  $\beta_0$  es la intersección de



la recta con el eje de ordenadas,  $\beta_i$  es la inclinación de la recta que forma la entrada con la salida,  $x_i$  es la variable independiente referente a la entrada,  $\epsilon_i$  es un término que forma parte de un vector de error, cada valor es distinto pero la media de este vector es 0 y  $n$  es el número de observaciones realizadas. En caso de que haya múltiples entradas y múltiples salidas, se usarán más variables independientes,  $x_i$  y  $x_j$  y si la respuesta constara de más de una valor también habría que añadir más valores,  $y_i$  y  $y_j$ .

### 3.3.2. SVR

Además de LR, se han usado otros regresores para poder comparar la eficacia de cada regresor. Es el caso de SVR, es un algoritmo que, al igual que LR, forma parte del SL y se conoce como el algoritmo usado para regresión de SVR, ya que este último permite tanto procesos de regresión como de clasificación y detección [42] [50] [51] [52]. Este regresor se basa en generar lo que se conoce como hiperplano, que correspondería con una línea que sería la aproximación de los puntos y un espacio alrededor de está línea,  $\epsilon$  - tube, que busca encontrar la mejor aproximación de todos los puntos intentando que la mayoría de estos queden en el interior de su espacio, como se observa en la Figura 3.7.

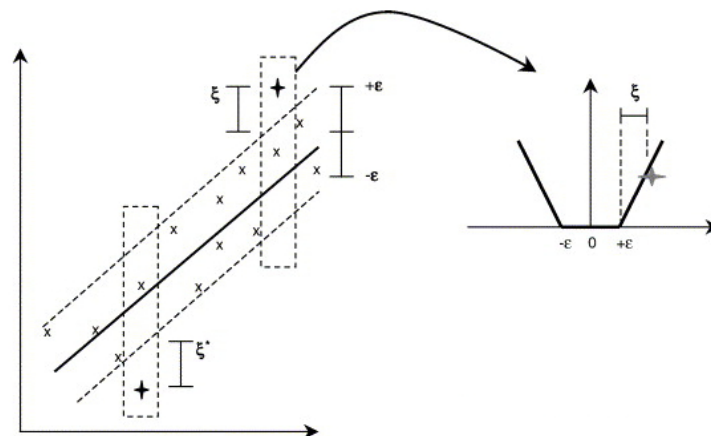


Figura 3.7: *Support Vector Regression* - Hiperplano [51].

SVR se entrena usando una función de pérdida simétrica, es decir, sigue una regla para penalizar igualmente los errores de estimación que no están en el hiperplano, tanto los que son al alza como a la baja de este hiperplano que se ha establecido. Una de las ventajas de SVR es que su complejidad computacional no depende de la dimensionalidad del espacio de entrada. Además, tiene una gran capacidad de generalización de los datos y es un modelo que ofrece una gran precisión [53].

Según la función de aproximación que se use y la precisión que se quiera conseguir el algoritmo variará. Para funciones que no sean lineales, se tendrá que usar lo que se conoce como espacio kernel, que permite más precisión y otorga un espacio con más dimensiones

permitiendo la aproximación de muchas funciones que no son lineales. Esto permite modelar relaciones complejas entre las distintas variables seleccionando una función kernel apropiada. Así, esta función transforma el espacio de entrada en un espacio de mayor dimensión en el que poder representar los datos en una forma lineal, como se presenta en la Figura 3.8. De tal forma que, dependiendo de los datos será mejor usar un kernel u otro para mapear los datos. La mejor forma para saber qué kernel es mejor es prueba y error y observar los resultados.

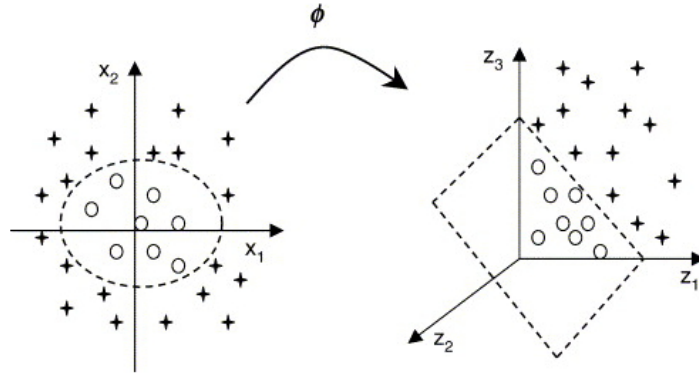


Figura 3.8: Transformación de los datos usando un kernel [51].

En la Ecuación (3.5), se encuentra la función que se usa a la hora de llevar a cabo SVR, siendo  $w$  la pendiente y  $b$  el desplazamiento de la función de regresión. Ambos parámetros se estiman minimizando la función de coste, ver la Ecuación (3.6), y varía según las condiciones que parecen en la Ecuación (3.7). También está el parámetro  $C$ , que es un coeficiente que determina el equilibrio entre el error y el componente de la Ecuación (3.6). Ésta corresponde a la función de pérdida utilizada para medir el riesgo entre el valor objetivo y el estimado. En esta misma función vemos el parámetro  $\varepsilon$ , que se utiliza para minimizar el riesgo nombrado anteriormente.

$$f(x) = wx + b \quad (3.5)$$

$$\min \frac{1}{2} \|w^2\| + C \sum_{i=1}^n L_{\varepsilon}(f(x_i), y_i) \quad (3.6)$$

$$L_{\varepsilon}(f(x_i), y_i) = \begin{cases} 0 & |y_i - f(x_i)| \leq \varepsilon \\ |y_i - f(x_i)| & \text{resto} \end{cases} \quad (3.7)$$

La Ecuación (3.5) no se adapta a todos los tipos de kernel que existen, para ello hay que usar la Ecuación (3.8). Siendo  $N$  el número de muestras,  $\alpha_i$  y  $\alpha_i^*$  son unos parámetros que regula cuanto va a afectar el kernel a esa muestra según lo acertada que

haya sido la estimación con respecto al valor verdadero, estos parámetros se conocen como los multiplicadores de Lagrange, y, por último,  $K$  es el kernel usado [54].

$$f(x) = \sum_{i=1}^N (\alpha_i - \alpha_i^*) K(x_i, x_j) + b \quad (3.8)$$

Existen distintos tipos de kernels y cada kernel tiene asociados una serie de funciones y algoritmos propios [55]. Uno de los que se ha usado es el lineal representado en la Ecuación (3.9). Este es el más simple de todos y, técnicamente, los datos no son proyectados en un espacio de mayor dimensión, por todo esto es el menos complejo de usar. Por otro lado, se ha utilizado también el kernel polinomial de la Ecuación (3.10), que si mapea los datos en un espacio de mayor dimensión. En este caso se introducen ya varios parámetros  $\gamma$ ,  $r$  y  $d$  y el grado  $d$  más usado suele ser 2, ya que para valores más grandes puede producirse un sobreajuste. Además, aparece el parámetro  $\sigma$ , que es el ancho del kernel, este parámetro afecta a como se distribuyen los datos en el espacio. Para terminar, está el Radial basis function kernel (kernel RBF), ver la Ecuación (3.11). Para este kernel, se pueden usar infinitas dimensiones, debido al incorporación de la exponencial.

$$K(x_i, x_j) = x_i^T x_j \quad (3.9)$$

$$K(x_i, x_j) = (\gamma \cdot x_i^T x_j + r)^d, \gamma > 0, r \geq 0, d \geq 2 \quad (3.10)$$

$$K(x_i, x_j) = e^{-\frac{\|x_i - x_j\|^2}{2\sigma^2}}, \sigma > 0 \quad (3.11)$$

### 3.4. Mínimos cuadrados

A lo largo del desarrollo del TFG, a parte de los regresores mencionados anteriormente se ha usado otro algoritmo para poder obtener los resultados necesarios. Es el caso de Least squares (LS), una técnica de regresión de análisis numérico en la que dados una variable independiente junto con su correspondiente variable dependiente y una ecuación se busca ajustar los distintos parámetros de la ecuación para obtener el mínimo error cuadrático [56] [57]. Es decir, se intenta minimizar las sumas de los cuadrados de las diferencias entre los valores reales y los resultados obtenidos con la función elegida. Esto se consigue a través de varias iteraciones que realiza el algoritmo internamente aplicando el residuo definido en la Ecuación (3.12). Siendo  $y_i$  los valores reales,  $x_i$  los valores de entrada,  $f(x_i)$  en el caso de este TFG es la ecuación de una circunferencia descrita en la Ecuación (3.13) y  $\beta_i$  el conjunto de valores que el algoritmo tiene que encontrar para

reducir el error. Además, este algoritmo es para la resolución de problemas de carácter no lineal como es el caso de una circunferencia, a diferencia de si fuera lineal que la función tendría que ser una recta.

$$\varepsilon = \sum_{i=1}^n (y_i - f(x_i, \beta_i))^2 \quad (3.12)$$

$$(x - a)^2 + (y - b)^2 = r^2 \quad (3.13)$$

# Capítulo 4

## Metodología

### 4.1. Introducción

En este capítulo, se explica el proceso que se ha llevado a cabo durante el desarrollo de este TFG. Además, se describe la información referente a los datos, tanto de dónde proceden como qué representan y el uso que se ha dado a los mismos. Por otro lado, se detalla las diferentes herramientas que se han usado para el tratamiento de estos datos. Todo el proceso se puede ver detallado en la Figura 4.1.

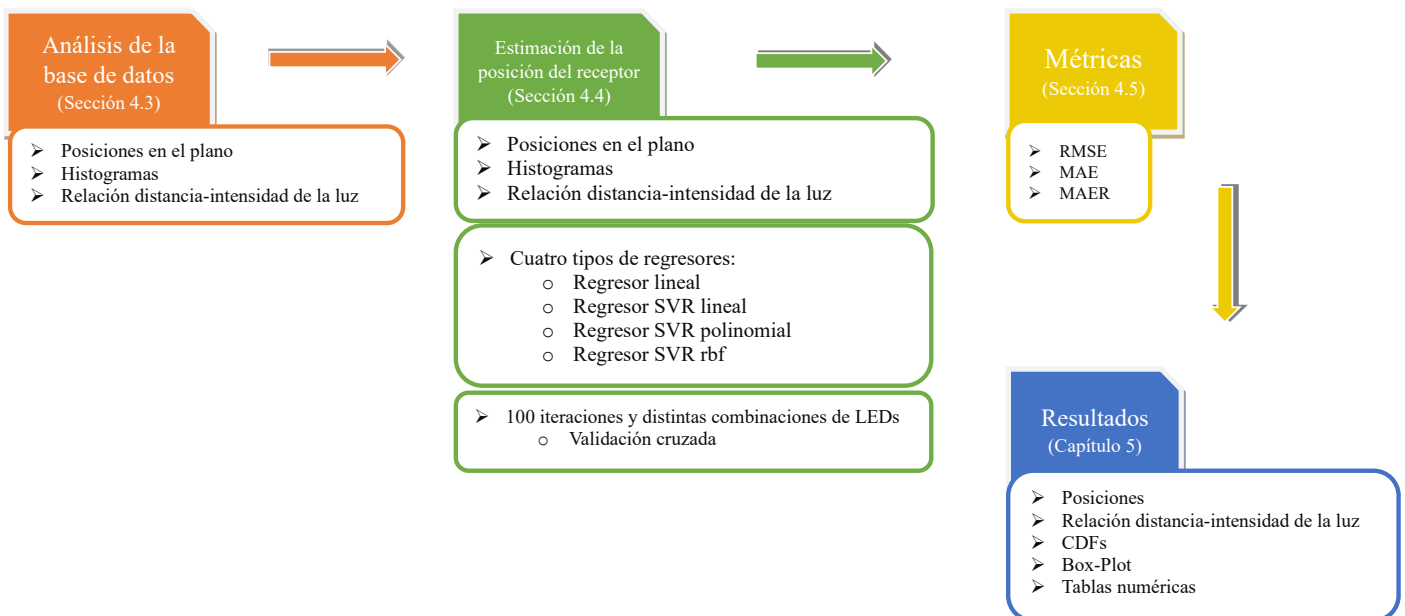


Figura 4.1: Esquema descriptivo de las pasos seguidos durante el TFG.

En este diagrama, se pueden ver todos los pasos que se han seguido una vez se ha

empezado a trabajar con la base de datos. Comienza con un primer análisis que incluye diversas representaciones gráficas como histogramas o la visualización de las posiciones en el plano. A continuación, se lleva a cabo la estimación de las coordenadas de las posiciones de con los distintos experimentos que se han hecho y los regresores usados. Tras la realización de los experimentos, se reportan métricas que evalúan los diferentes algoritmos ML. Por último, se presentan los resultados calculados, las gráficas y las tablas.

## 4.2. Base de datos

En esta sección, se explicará los valores que conforman la base de datos y se ofrecerá información sobre la misma. Cabe destacar que la toma de datos y la formación de la base de datos no han sido parte de este TFG, sino que la base de datos se ha tomado del artículo [58]<sup>1</sup>. La formación de la misma se ha llevado a cabo durante la realización de la investigación [59].

La base de datos esta formada por puntos espaciales contenidos en una rejilla 3m x 3m, en el que se disponen 4 LEDs a una altura de 1,284m del área de test, como se puede ver en la Figura 4.2.



Figura 4.2: Escenario usado en la toma de datos [59].

Estos LEDs emiten cada uno luz a una frecuencia distinta, las cuales son 500, 1000, 2000 y 4000Hz. En el plano base, se ha colocado un receptor que consta de un fotodiodo. Por lo cual, en cada posición el fotodiodo recibirá más intensidad de una frecuencia que de otra según su posición en el espacio. Se han medido en un total de 128 posiciones las intensidades de la luz que llegan procedentes de los LEDs. En la Figura 4.3, se tiene una representación de las posiciones medidas en 2D del escenario.

<sup>1</sup><https://iee-dataport.org/documents/machine-learning-rss-based-visible-light-positioning>

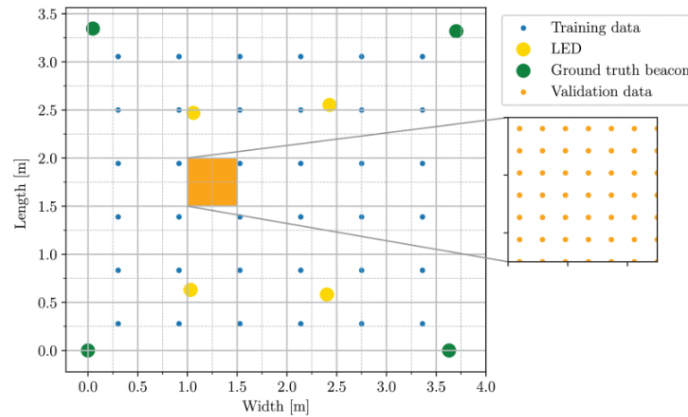


Figura 4.3: Escenario con la posición de cada LED [59].

Así, se fue completando la base de datos. Por un lado, con las coordenadas  $(x,y)$  reales de cada posición en el momento de la medida de las intensidades, que consiste en una matriz con un tamaño de  $(128, 2)$ . Por otro lado, una matriz con la intensidad referente a cada uno de los cuatro LEDs captada por el fotodiodo, obteniéndose una matriz de tamaño  $(128, 4)$ .

### 4.3. Análisis de la base de datos

A lo largo de esta sección, se va a exponer el análisis que se ha llevado a cabo de la base de datos. En primer lugar, aclarar que todo el TFG se ha llevado a cabo mediante la programación de *scripts*, todos ellos en el lenguaje de programación *Python*. Para ello, se han dividido los procesos de análisis y tratamiento de los datos en seis *scripts* distintos. Los *scripts* se pueden encontrar en el repositorio de GitHub accesible desde el enlace <https://github.com/jimenezmr/TFG>. El primero de ellos se encarga de realizar un análisis de la base de datos, es decir, de cómo son esos datos con los que se va a trabajar para posteriormente hacer una representación de los mismos y poder analizarlos detalladamente de forma gráfica. Para ello, se ha desarrollado un archivo de *Python*, en el que a través de diferentes cálculos y gráficas se ha obtenido información de los datos. Además, se han utilizado diferentes clases ya definidas con sus métodos que se irán nombrando todas ellas a lo largo del documento. Una de las que más se ha usado a lo largo de todo el desarrollo es la clase *numpy*, que sirve para trabajar con arrays de datos. Los siguientes cuatro *scripts* desarrollan cada uno de los cuatro regresores usados durante el TFG y se detallarán más adelante. Por último, un sexto *script* que contiene el código referente a los distintos errores que se han usado para las mediciones y que se ejecutará cuando se ejecute el código de cada regresor. Esto es debido a que como se emplean los mismos errores para analizar todos los regresores, lo más eficiente consistía en desarrollar un *script*, que pudiera usarse para todos los casos de regresores.

En primer lugar, se ha realizado una representación 2D del escenario. Es decir, se han representado cada uno de los LEDs en su ubicación, así como cada una de las posiciones que estaban en la base de datos, esto se puede ver en la Figura 4.4. Esto se ha llevado a cabo para tener una imagen clara de los datos y como se distribuían a lo largo del espacio. En la Figura 4.4, están representadas como puntos azules cada una de las posiciones y como cruces de distintos colores cada uno de los LEDs. Igualmente las unidades de la gráfica son metros. Para llevar a cabo esta representación se ha usado, además, la clase *matplotlib.pyplot*.

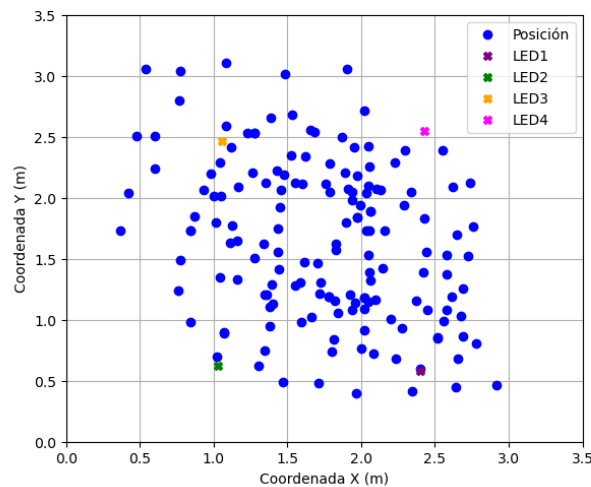


Figura 4.4: Distribución de la base de datos sobre el plano XY.

A continuación, se han creado diferentes gráficas para poder comparar los datos de forma visual. Una de estas representaciones se puede ver en la Figura 4.5 y consiste en una gráfica 3D, donde se representa la intensidad en dB de cada LED para cada una de las 158 posiciones en las que se ha situado el receptor, siendo cada color la intensidad de cada LED, es decir, el azul son las intensidades correspondientes a cada posición provenientes del LED 1, el naranja para el LED 2, el verde para el LED 3 y el rojo para LED 4. Para esta representación en 3D se ha utilizado la clase *mpl\_toolkits.mplot3d*. Estas cuatro gráficas se encuentran representadas en el mismo plano en el anexo en la Figura A.1. Para tener una visión más clara de estos datos, también se ha realizado esta representación separando cada LED en una gráfica distinta con las 158 posiciones, obteniendo un total de cuatro gráficas. Esta representación se puede ver en la Figura 4.6. Continuando con el análisis, para tener más información de los datos se ha realizado un histograma de la intensidad recibida de los LEDs, como se puede observar en la Figura 4.7. Se han establecido distintos rangos para la intensidad y conjuntamente se ha separado la intensidad de cada LED por un color.



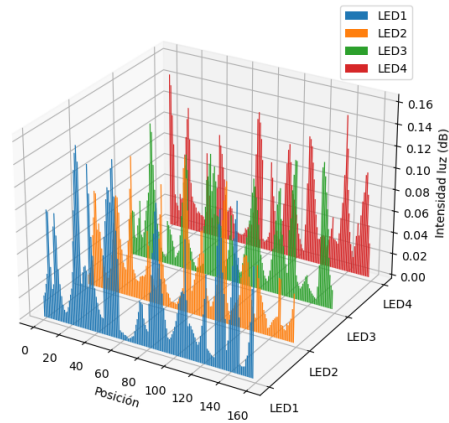


Figura 4.5: Intensidad recibida de cada LED en cada posición separadas por LEDs por colores.

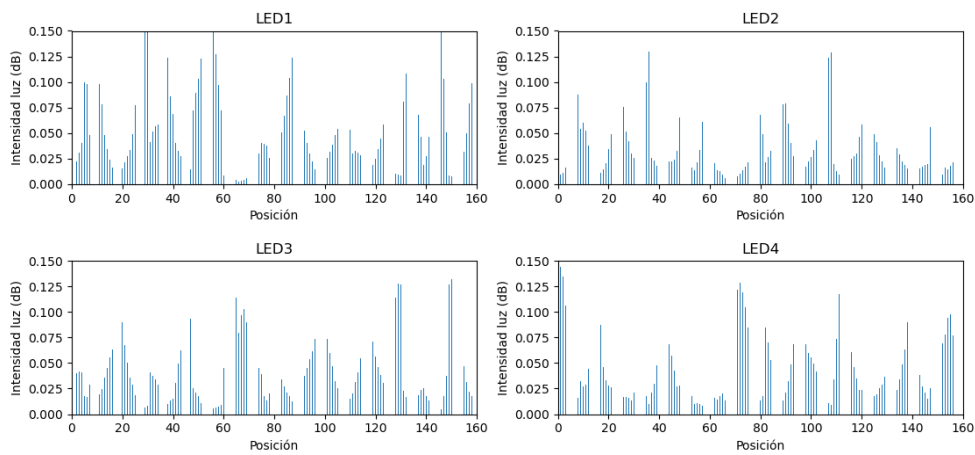


Figura 4.6: Intensidad recibida de cada LED en cada posición separadas por LEDs en distintas gráficas.

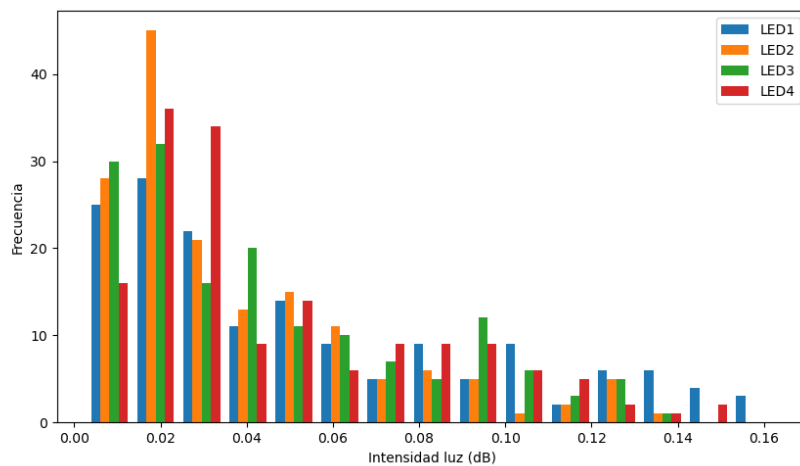


Figura 4.7: Histograma Intensidad-Frecuencia separados por colores por LEDs.

Viendo los resultados se observa que la mayoría de valores de la intensidad se encuentran para valores menores a 0.04dB. Esto concuerda con el resto de gráficas, ya que la mayoría de posiciones reciben la mayor intensidad de luz si justo se encuentran debajo o muy cerca de un LED y, entonces, reciben poca luz del resto. Esto se observa por ejemplo en la posición 128 (1,12m, 2,42m) donde el LED 3 tiene 0.127dB, pero el resto de LEDs tienen una intensidad de luz inferior a 0.037dB, ya que esa posición coincide justo debajo del LED 3. Además, si no están cerca de ningún LED, la luz que recibirán de todos ellos será escasa y similar. Es el caso de la posición 91 (1,83m, 1,58m) en la que la intensidad de luz recibida de los LEDs no supera en ningún caso los 0.053dB.

Con el fin de tener una mejor visión del histograma de los datos, se ha realizado otra representación de los datos, que se puede ver en la Figura 4.8. Para ello, se han distribuido los datos del anterior histograma en cuatro histogramas distintos separados según el LED. Conjuntamente se ha introducido la curva de aproximación de cada uno de ellos de color naranja para ver la evolución según va aumentando los valores de intensidad. Lo que se puede observar en todas las gráficas es que la gran mayoría de valores se encuentran en los valores cercanos a cero, y a medida que aumenta la intensidad va disminuyendo el número de valores para esa intensidad, correspondiéndose así con una distribución monótona decreciente. Esto concuerda con las posiciones, ya que a mayor cercanía a un LED, mayor es la intensidad que se recibe del mismo. Pero esta mayor intensidad solo se puede obtener si se encuentra cerca de un LED, lo que requiere que esté lejos del resto y, por tanto, reciba menos intensidad de ellos.

Por otro lado, la curva de aproximación que se representa es asimétrica y en todos los casos consta de rizado, siendo el caso del LED 1 el que menos rizado tiene. Aunque no es un rizado muy pronunciado, demostrando que el valor escogido para la curva de aproximación es correcto. La curva de aproximación se ha conseguido mediante la obtención de un polinomio de grado 9 utilizando el método *poly1d* de la clase *numpy*. Ya que sino fuera el correcto, en la aproximación se obtendría rizado, por lo que se estaría sobreestimando. Y en el caso contrario, si la curva fuera demasiado suave se estarían perdiendo datos y la curva de aproximación no sería representativa del histograma. También en cada gráfica se observa una línea punteada de color negro, ésta se corresponde con el percentil 50 de la intensidad de la luz de todas las muestras. En todas las gráficas se encuentra desplazada a la derecha del rango medio del histograma. Esto concuerda con los datos obtenidos en el histograma. Ya que la mayoría de valores se encuentran para los valores de menor intensidad.

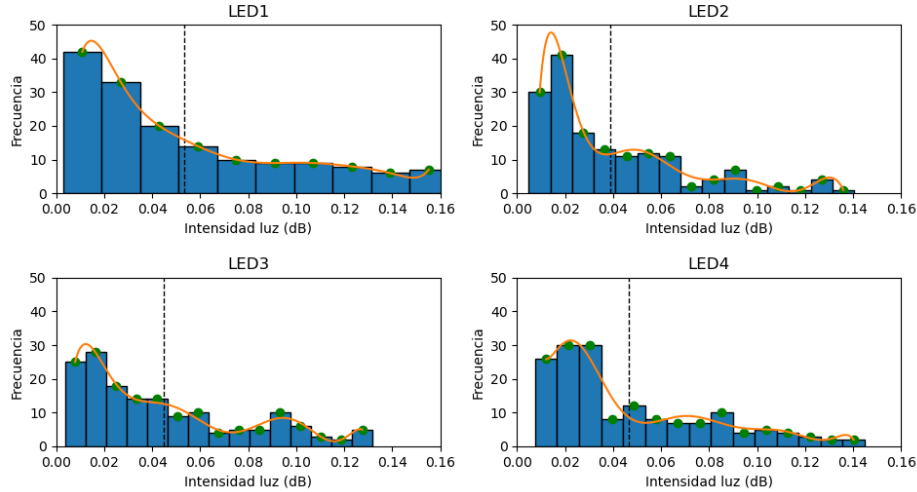


Figura 4.8: Histograma Intensidad-Frecuencia y curva de aproximación separados por gráficas por LEDs.

Finalmente, para mostrar la relación que existe entre la distancia a un LED y la intensidad obtenida se ha calculado la distancia de cada posición a cada LED, se ha representado cómo afecta esta distancia a la intensidad, esta relación va definida en las Ecuaciones 4.1 y 4.2 [59]. En la Ecuación 4.1, se calcula la señal transmitida por un LED. En la cual,  $A_r$  es el área efectiva de un fotodiodo,  $d_q$  la distancia estimada entre un LED y el fotodiodo,  $G_q$  la distribución angular de la potencia de salida,  $\theta_q$  el ángulo de inclinación,  $\psi_q$  el ángulo de acimut y  $\gamma_q$  es el ángulo de incidencia en el fotodiodo. Por otro lado, la Ecuación 4.2 se corresponde con la intensidad de la señal en el receptor. En la cual,  $\alpha_q$  es el valor calculado anteriormente en la Ecuación 4.1,  $R_p$  la responsividad del fotodiodo,  $s_q$  la señal transmitida por un LED,  $\beta$  el ángulo de inclinación y  $n$  el componente de ruido, que es modelado como ruido gaussiano.

$$\alpha_q = \frac{A_r}{d_q^2} G_q(\theta_q, \psi_q) \cos \gamma_q \quad (4.1)$$

$$r(t) = \sum_{q=0}^3 \alpha_q R_p s_q(t) + \beta + n(t) \quad (4.2)$$

En este TFG, para obtener la distancia entre una posición y un LED se ha usado la fórmula de la distancia euclídea que se basa en el teorema de Pitágoras. Esta fórmula se puede ver en la Ecuación (4.3), en la que  $x_p$  e  $y_p$  son las coordenadas  $x$  e  $y$  respectivamente de una posición donde se ha situado el fotodiodo y  $x_L$  e  $y_L$  son las coordenadas  $x$  e  $y$  de un LED. Se puede ver el esquema de la distancia calculada en la Figura 4.9.

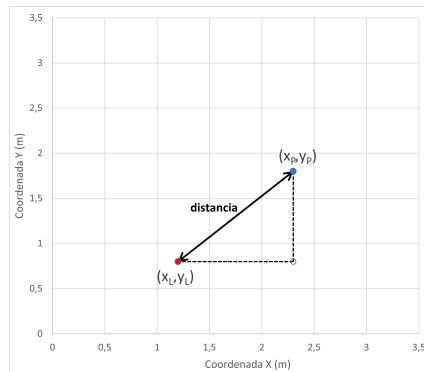


Figura 4.9: Gráfico de como se cálcula la distancia entre un LED y una posición.

$$dist = \sqrt{(x_P - x_L)^2 + (y_P - y_L)^2} \quad (4.3)$$

Así, se puede observar en la Figura 4.10, para cada LED y según la distancia, cuál es el valor de la intensidad en cada posición y cómo evoluciona esa intensidad mediante una curva de aproximación, obtenida por aproximación con un polinomio de grado 5. En azul, se representan cada una de las posiciones en las que se ha situado el receptor y en rojo la curva de aproximación aplicada a las muestras. Como se puede ver en la curva de aproximación no hay rizado y sigue la forma de una función decreciente uniforme. Concluyendo, cuanto menor es la distancia al LED mayor es la intensidad y a medida que esa distancia aumenta disminuye la intensidad. Esta afirmación concuerda con los datos obtenidos en el histograma.

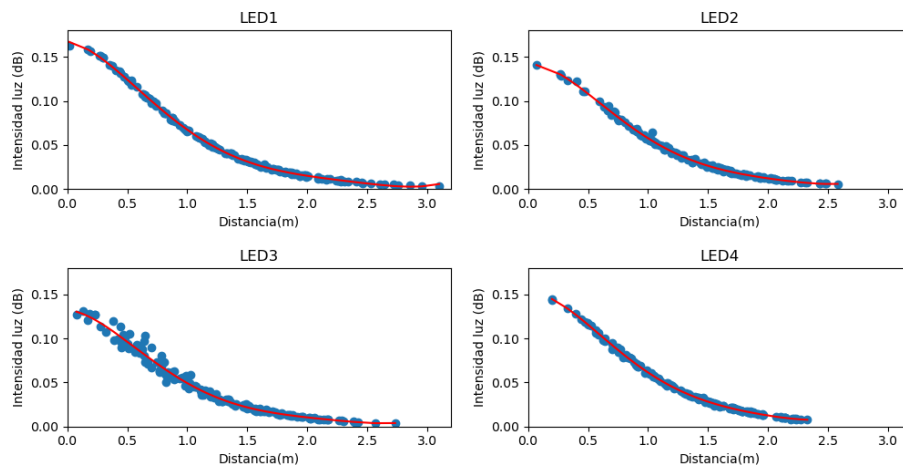


Figura 4.10: Relación Intensidad-Distancia separadas por gráficas por LEDs.

## 4.4. Estimación de la posición del receptor

A parte de un análisis de la base de datos, se han llevado a cabo diferentes desarrollos para comparar, en este caso, cuatro regresores y definir cual de ellos es más preciso. En esta sección, se va a explicar el proceso que se ha seguido para poder obtener los datos de los distintos regresores.

En primer lugar, se mencionará cada uno de los cuatro regresores que se han elegido para este TFG. Estos regresores son el regresor lineal, regresor *SVR* lineal, regresor *SVR* polinomial y regresor *SVR* RBF. Para cada uno de ellos se ha desarrollado un *script*, en el que llevar a cabo el experimento y la estimación, así como el estudio de los resultados. Además, dado que todos siguen un mismo proceso se ha creado un quinto *script* para el cálculo de errores, al que todos llaman en cierto momento para obtener esos errores y, así, hacer un desarrollo más eficiente.

De toda la base de datos se han escogido de forma aleatoria mediante el uso de la clase *random* 126 muestras que se usarán para el grupo de entrenamiento. Quedando 32 muestras para el grupo de test. Por lo que en el caso de este TFG, del total de muestras, casi el 80 % (79,75 %) se ha usado para el entrenamiento y aproximadamente el 20 % (20,25 %) se ha utilizado en el grupo de test. De todas las posiciones de la base de datos había tres en concreto, la número 29, 77 y 129, que, dado su lugar en el plano, alteraban en gran medida los datos de forma perjudicial, por lo que se han introducido de forma obligatoria en el grupo de entrenamiento. En concreto, sus coordenadas son: la número 29 con la posición en (2.4m, 0.6m), la número 77 con la posición en (1.02m, 0.7m) y la número 129 con la posición en (1.12m, 2.42m). El problema con su posición era que estaban muy cerca de algún LED, por lo que los regresores al intentar estimar su posición no sabían donde ubicarlo. A continuación de la obtención de los dos grupos, se han entrenado los regresores mediante el grupo de test. De tal forma, que la intensidad de la luz era la variable independiente. En el caso de la variable dependiente, existen distintos casos que son los siguientes.

- Experimento tipo 1: En este caso, se usan dos regresores en el experimento para obtener dos variables dependientes. Una de las variables dependientes es la coordenada  $x$  de la posición y la otra variable dependiente corresponde a la coordenada  $y$ .
- Experimento tipo 2: Para este caso, se utilizan cuatro regresores para las variables dependientes. Obteniendo de cada regresor una distancia a cada LED. Así, se usa un regresor para cada uno de los LEDs. Además, en el caso de la estimación de las distancias, después de la estimación, mediante la ecuación de la circunferencia, como se puede ver en la Ecuación (4.4) y la intersección de circunferencias se ha

obtenido las coordenadas de las posiciones. Para ello se ha utilizado el algoritmo *LS*, que se puede encontrar explicado en mayor detalle en la Sección 3.4. Esto solo ha sido posible para las combinaciones en las que se tenían en cuenta tres o más LEDs, ya que sino no se obtenía una única posición. Esto se ha resuelto usando la clase *scipy.optimize*.

$$(x - a)^2 + (y - b)^2 = r^2 \quad (4.4)$$

También para poder observar la diferencia según el número de LEDs y ver la relevancia de cada uno de ellos a la hora de una estimación más precisa, se ha desarrollado otro análisis, que se ejecuta en el *script* compartido por todos los regresores. En él se ejecutan los dos experimentos anteriormente mencionados teniendo en cuenta todas las posibles combinaciones de LEDs. Es decir, primero solo teniendo en cuenta la intensidad del primer LED, después solo la del segundo y así sucesivamente. Más tarde, solo se tienen en cuenta los dos primeros LEDs, posteriormente solo el primero y el tercero, etc. De esta forma se obtienen todas las combinaciones que pueden existir con los LED. Estas combinaciones son: LEDs 1, LEDs 2, LEDs 3, LEDs 4, LEDs 1-2, LEDs 1-3, LEDs 1-4, LEDs 2-3, LEDs 2-4, LEDs 3-4, LEDs 1-2-3, LEDs 1-2-4, LEDs 1-3-4, LEDs 2-3-4 y LEDs 1-2-3-4. Este proceso, como se ha mencionado, se realiza tanto obteniendo las coordenadas como la distancia a cada LED de los que se ha escogido en esa combinación.

Todo este proceso se ha ejecutado para cada uno de los cuatro regresores variando la clase importada para el uso del regresor en cuestión. Siendo la clase usada para el regresor lineal *sklearn.linear\_model*. En cambio, el resto de regresores comparten la clase usada, *sklearn.svm*, pero varían el kernel usado ya que cada uno dispone de uno distinto siendo estos *linear*, *poly* y *rbf*.

Posteriormente, se ha llevado a cabo un análisis de la estimación a través de diferentes parámetros y gráficas que se expondrán en el siguiente capítulo.

## 4.5. Métricas

En esta sección, se van a exponer las distintas métricas que se han utilizado para la medida y evaluación del rendimiento de cada regresor y tipo de experimento, ya que todos los experimentos deben ser evaluados con el mismo criterio para que los resultados se puedan considerar válidos.

En primer lugar, es importante mencionar que cada experimento en cada regresor se ha realizado un total de 100 veces para cada combinación de LEDs. Para lo cual, en cada iteración se han elegido de forma aleatoria las muestras utilizando la clase *random*,

siempre siguiendo el proceso comentado en la sección anterior. Esto se ha realizado para poder asegurarse que los resultados no dependen de las muestras que se escojan, por lo tanto, se ha empleado la técnica de validación cruzada explicada en la Sección 3.4.

Una vez se han realizado las 100 iteraciones para todas las combinaciones y se han recogido las estimaciones de los resultados, se procede a realizar los cálculos para poder analizar el rendimiento de los regresores. Estas métricas están explicadas a continuación.

#### 4.5.1. Error cuadrático medio (RMSE)

En primer lugar, se ha calculado el RMSE. Para ello, una vez se han realizado las estimaciones se han seguido una serie de pasos que serían los siguientes.

1. Para cada posición de las que se ha calculado su coordenada, se resta el valor real ( $x_R$ ) menos la estimación ( $x_E$ ) y se eleva al cuadrado.
2. Se calcula la media del conjunto de valores del array y se hace la raíz cuadrada, como se puede ver en la Ecuación (4.5). Como hay 100 iteraciones por cada iteración se creará un valor que estará guardado en un array.
3. Se realiza la media de los valores del array que se han obtenido en cada iteración:

$$RMSE = \sqrt{\frac{\sum_{i=0}^{N-1} (x_i - \hat{x}_i)^2}{n}} \quad (4.5)$$

#### 4.5.2. Error medio absoluto (MAE)

En segundo lugar, se ha calculado el MAE. Para lo cual se han seguido los siguientes pasos.

1. Tras las estimaciones se restan los valores reales y los estimados para cada posición y nos quedamos con el valor absoluto. Obteniendo un array para cada coordenada.
2. Se hace la media de la diferencia entre todas las posiciones, como se puede ver en las Ecuaciones (4.6). Pero como se realizan 100 iteraciones, se obtiene un array con una media por iteración.
3. Se realiza la media de entre todos los valores de las distintas iteraciones para obtener el MAE:

$$MAE = \frac{\sum_{i=0}^{N-1} |x_{Ri} - x_{Ei}|}{n} \quad (4.6)$$

### 4.5.3. Error medio relativo (MAER)

Finalmente, el último error que se ha calculado es el MAER. A continuación, se presentan los pasos que se han seguido para calcularlo.

1. Una vez se tienen las estimaciones, se restan los valores de las coordenadas reales menos las estimadas ( $x_E$ ) y ese valor se divide entre el valor de la coordenada real. De este valor nos quedamos con su valor absoluto.
2. Como lo que se tiene es una array con los valores que se han calculado de cada posición, se hace la media para tener un valor concreto, como se ve en la Ecuación (4.7). Pero al haber las 100 iteraciones se obtiene otro array, con un valor por iteración.
3. Para obtener el valor final, se hace la media del valor de cada iteración:

$$MAER = \frac{\sum_{i=0}^{N-1} \left| \frac{x_{Ri} - x_{Ei}}{x_{Ri}} \right|}{n} \quad (4.7)$$



# Capítulo 5

## Resultados

### 5.1. Introducción

Una vez que se han llevado a cabo cada uno de los experimentos, los datos obtenidos se han representado mediante diferentes gráficas y se han introducido en una tabla los resultados de estos experimentos.

A continuación, se va a presentar cada experimento. En primer lugar, se han colocado las gráficas del experimento en cuestión para posteriormente tener en distintas tablas con números los datos que se han obtenido. Cabe destacar que no todas las gráficas y las tablas se encuentran en esta sección, algunas de ellas se han localizado en la sección de Anexo A.2, debido a la cantidad de las mismas. En el siguiente apartado solo se representan las gráficas en las que se tiene en cuenta la información de todos los LEDs, ya que aunque los experimentos se han realizado para diferentes combinaciones, los resultados más representativos se obtienen en las gráficas con la combinación de los cuatro LEDs.

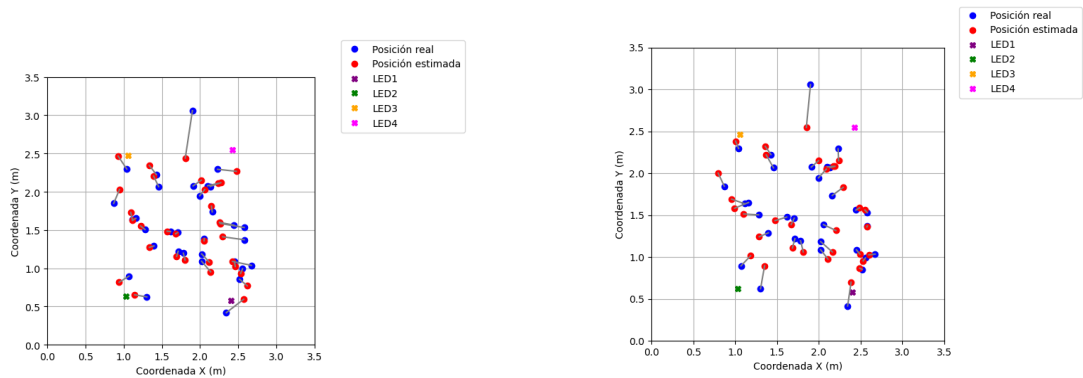
### 5.2. Resultados

#### 5.2.1. Regresor lineal

Como se ha comentado a lo largo del TFG, para cada regresor se han llevado a cabo dos tipos de experimento: estimando la posición (experimento tipo 1) y estimando la distancia a cada LED (experimento tipo 2). En primer lugar, vamos a comparar los resultados estimados de las posiciones del receptor en el plano frente a las posiciones reales conocidas. Para ello, se han configurado dos regresores, uno para la coordenada  $X$  y otro para la coordenada  $Y$ . De tal forma, que se han obtenido las posiciones del receptor utilizando el experimento tipo 1 y el tipo 2, mostrándose los resultados en las Figuras 5.1a y 5.1b

respectivamente. En azul se representan las posiciones reales y en rojo las estimaciones, además cada par de posiciones está unido por una línea para ver la relación entre ellas. En las figuras, a parte de los puntos rojos y azules, también se ven cuatro cruces que se corresponden con cada uno de los LEDs. Como se puede observar en las posiciones centrales (es decir, las que están entre los cuatro LEDs) tienen una menor variación y, a medida que el receptor se acerca a un extremo del área de test también es mayor la variación de la estimación de su posición estimada con respecto a la real).

Este último, lo encontramos representado en la Figura 5.1b usando el regresor lineal. En este caso, se han usado cuatro regresores uno para cada LED. Comparando estos resultados con los de la estimación de las posiciones, se obtienen unos resultados muy similares, destacando, que en el centro del plano es donde el experimento tiene mayor precisión. Esto se debe a que es donde más información se tiene de los cuatro LEDs, y al ser este el resultado de la combinación en la estimación de cuatro regresores esas posiciones se benefician de esto.



(a) Distribución de los datos reales y los estimados sobre el plano con el experimento tipo 1. (b) Distribución de los datos reales y los estimados sobre el plano con el experimento tipo 2.

Figura 5.1: Box-plot del RMSE para el regresor lineal.

También se ha representado la relación que existe entre la distancia de las posiciones con los LEDs y las intensidades separando las distancias reales de las estimaciones en dos gráficas. Así, se tiene por un lado las distancias a los LEDs en cada posición (eje de ordenadas) y, por otro, la intensidad de luz recibida en cada posición (eje de abscisas). Con estos valores se ha calculado la curva de aproximación para poder observar la tendencia de los resultados. De tal forma que en la Figura 5.2, se puede ver perfectamente estas relaciones para el caso del LED 1 para el experimento tipo 1, que se ha escogido a modo representativo de todos los LEDs. Las gráficas de los resultados para el resto de LEDs están, como se ha mencionado antes, en la sección de Anexo A.2 y en ellas se obtienen resultados muy similares a los obtenidos para el LED 1. Si nos fijamos en la curva de aproximación que se encuentra representada en rojo, se ve que ambas relaciones tiene forma de parábola y pasan prácticamente por los mismos valores con la diferencia de que

la estimada ha eliminado los valores de mayor distancia.

En el caso del experimento tipo 2, la gráfica con la relación entre distancia e intensidad se puede ver en la Figura 5.3. Aquí, al igual que antes la forma es de una parábola y también se observa que las distancias de mayor valor se eliminan. Otra observación a tener en cuenta que ocurre en las dos gráficas de la estimación es que los puntos están más dispersos y no se corresponden de forma tan precisa con su curva de aproximación como si lo hacen sus respectivas reales.

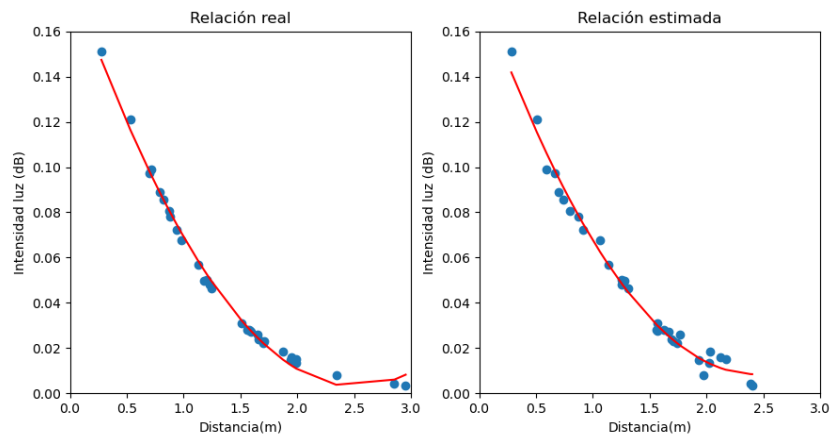


Figura 5.2: Relación Intensidad-Distancia de los datos reales y los estimados para el LED 1 con el experimento tipo 1.

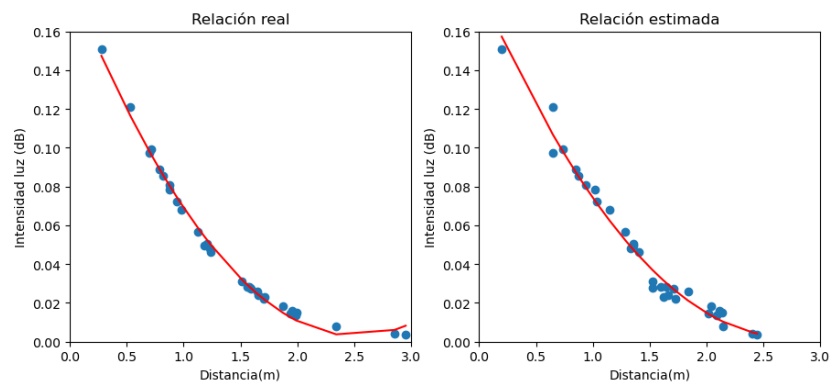


Figura 5.3: Relación Intensidad-Distancia de los datos reales y los estimados para el LED 1 con el experimento tipo 2.

Siguiendo con otras representaciones, se encuentra la Cumulative Distribution Function (CDF) para ver cómo se distribuyen los datos y como va cambiando el RMSE según varían los datos que se puede ver en la Figura 5.4. Además, también se han representado los percentiles 50 y 90, que se corresponden con las líneas horizontales punteadas. En primer lugar, se ha representado la CDF de las coordenadas tanto para el experimento tipo 1 como el tipo 2. Como se puede ver, el experimento tipo 2 obtiene un menor RMSE

que el tipo 1. También se observa que en el caso del experimento tipo 1, la coordenada  $X$  obtiene un mejor resultado, pero en el caso del tipo 2 la que obtiene unos resultados más precisos es la coordenada  $Y$ . Además, se puede observar, entre la mejor estimación, la coordenada  $X$  obtenida de la distancia, y la peor, la coordenada  $X$  siendo la estimación de la coordenadas, esta última prácticamente dobla el valor de la primera. Estando los valores entre el 0.11m y 0.18m para el percentil 50 y 0.125m y 0.225m para el percentil 90.

Para tener la comparación entre las distancias se ha representado la CDF de la estimación de las distancias separadas por LEDs que se encuentra en la Figura 5.5. Se puede ver que entre los propios LEDs hay poca variación entre sus RMSE, siendo el peor el LEDs 1 y el mejor el LED 4. Esto se puede deber a que las muestras escogidas estaban más alejadas del LEDs 1 y más cerca del LED 4 mejorando la estimación para uno y empeorando la del otro. En general, la CDF de los LEDs tiene un parecido con la CDF anterior de la coordenada y obtenida la estimación de las distancias. En cuanto a lo percentiles, se encuentran en torno al para el de 50 y entre el 0.15m y el 0.2m para el de 90.

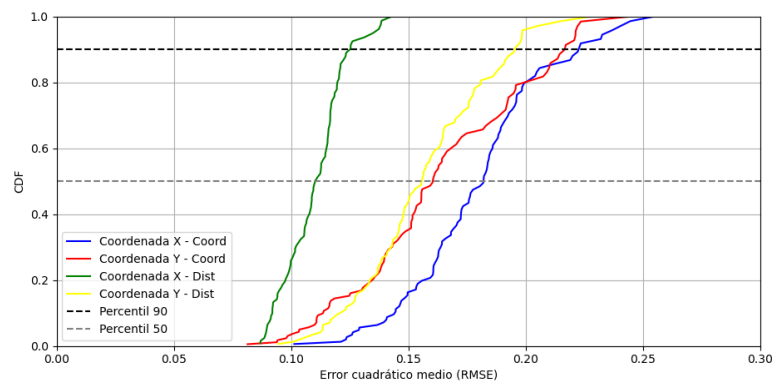


Figura 5.4: CDF del RMSE de las coordenadas ( $x,y$ ) obtenidas para el experimento tipo 1 (coord.) y el experimento tipo 2 (dist.).

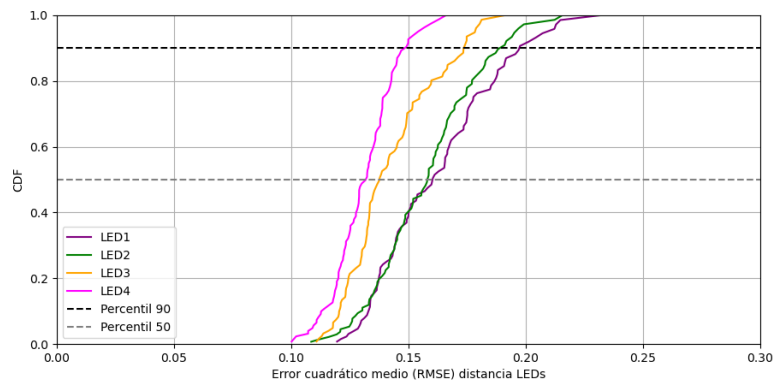
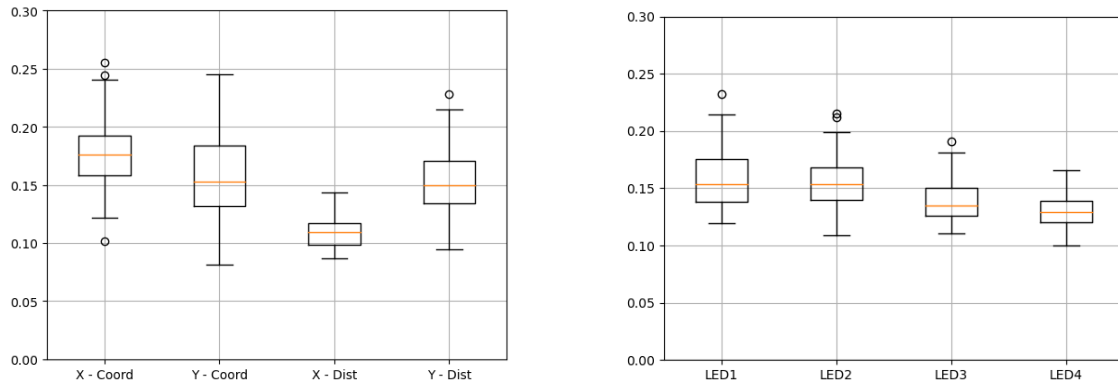


Figura 5.5: CDF del RMSE de las distancias obtenidas para el experimento tipo 2.

Por último, en cuanto a representaciones se ha elegido el diagrama de cajas del RMSE para tener otra representación de los datos y poder comparar mejor ambos experimentos. Los resultados de ambos tipos de experimentos están en la Figura 5.6a. Lo que se observa corrobora perfectamente lo que se había visto en las representaciones de la CDF. Los menores valores de RMSE son para la coordenada  $X$  de la distancia y los peores los de la coordenada  $X$  de las coordenadas. Pero en este caso, podemos destacar también que las coordenadas  $X$  en este caso obtienen unos valores más concretos y menos dispersos entre ellos que la coordenada  $Y$  como se puede observar por el ancho de las cajas. También destacar que la coordenada  $X$  de las coordenadas tiene más valores atípicos que el resto viendo los círculos que se encuentran fuera de su rango. Además, también se observan los rangos y el percentil 50 que coinciden con los mencionados en la CDF.

A parte de la anterior representación, se ha decidido para el experimento tipo 2 ver la diferencia entre la distancia estimada a cada LED con un diagrama de cajas para tener la comparación, ésta es la Figura 5.6b. Así, se visualiza que todos obtienen unos resultados similares. Repitiéndose la conclusión de que el de menor RMSE es el LED 4 y el de mayor RMSE es el LED 1. Otro punto a comentar es que la mayor parte de ellos tienen los mismo valores de RMSE en su clasificación, a diferencia del caso de las coordenadas que eran más dispersos y el ancho de las cajas era mayor.



(a) Box-plot del RMSE para el experimento tipo 1.

(b) Box-plot del RMSE para el experimento tipo 2.

Figura 5.6: Box-plot del RMSE para el regresor lineal.

Una vez comentadas todas las representaciones gráficas de los errores, se procede a hablar de los errores en sí. Para ello, entre todos las métricas se ha escogido los valores de RMSE para las coordenadas estimadas, para las distancias estimadas y para las coordenadas calculadas de estas últimas estimaciones. Además, se han representado los errores para las combinaciones de 3 o más LEDs. Estas tres tablas se pueden ver en las Tablas 5.1, 5.2 y 5.3. Además, estos valores son los obtenidos después de hacer la media entre 100 realizaciones.

De esta modo, se pueden observar cómo afecta el realizar un método u otro y la importancia de cada LED a la hora de ver la eficacia de los regresores. Así, se observa que los menores valores de RMSE se obtienen con el uso de los cuatro LEDs, menos en el caso del LED 2 en la estimación de las distancias, Tabla 5.3, en el que la combinación LED 1-2-3 obtiene mejores resultados que los cuatro LEDs juntos. Esto se debe a la relevancia que tiene cada LED a la hora de calcular las posiciones o los valores de RMSE. Otro dato a mencionar es que la combinación de LEDs 1-2-4 es la que obtiene los menores valores de RMSE. Del mismo modo, la peor combinación es la de LEDs 1-2-3. Por último, recalcar algo mencionado anteriormente y es que la estimación de la posición del receptor utilizando el experimento tipo 2 ha obtenido valores menores de RMSE que la estimación de la posición del receptor utilizando el experimento tipo 1. Asimismo, el cálculo de las coordenadas obtenido de la estimación de las distancias es el que tiene el menor RMSE, por lo que se puede asumir que es la forma más precisa usando el regresor lineal de obtener las coordenadas de una posición, obteniendo errores inferiores a 0.23m y 0.21m en las coordenadas  $x$  e  $y$ , respectivamente. Para los MAE y MAER, los valores menores se encuentran en el experimento tipo 2 con valores entre 0.15m y 0.09m y 0.07m y 0.17m, respectivamente. Estos datos están en el anexo en las Tablas A.15 y A.18.

Tabla 5.1: Error cuadrático medio para el regresor lineal en el caso del experimento tipo 1 para cada combinación de LEDs.

RMSE – Coord.	LEDs				
	123	124	134	234	1234
Coord. X (m)	0,255±0,062	0,191±0,034	0,193±0,035	0,246±0,055	0,177±0,029
Coord. Y (m)	0,156±0,028	0,167±0,028	0,2±0,043	0,212±0,041	0,15±0,032

Tabla 5.2: Error cuadrático medio para el regresor lineal en el experimento tipo 2 para obtener unas coordenadas para cada combinación de LEDs.

RMSE – Dist (coord)	LEDs				
	123	124	134	234	1234
Coord. X (m)	0,222±0,067	0,130±0,019	0,155±0,030	0,117±0,014	0,108±0,014
Coord. Y (m)	0,167±0,026	0,160±0,019	0,201±0,042	0,171±0,026	0,155±0,024

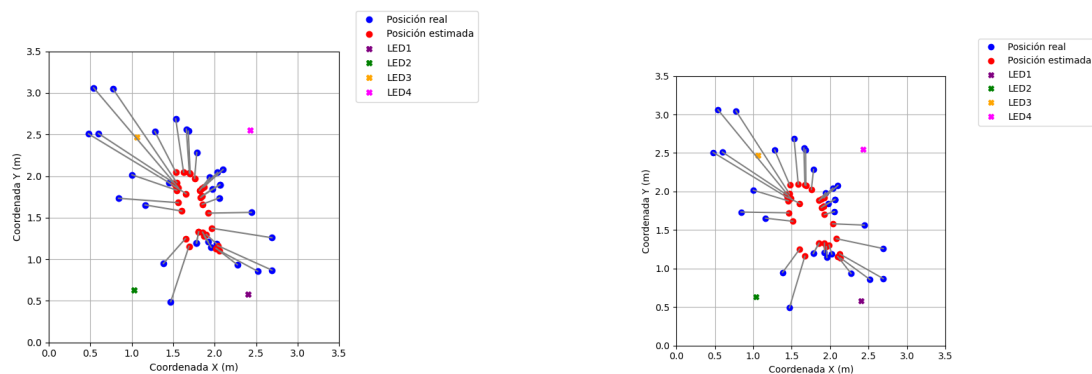
Tabla 5.3: Error cuadrático medio para el regresor lineal en el caso de estimar las distancias para cada combinación de LEDs.

RMSE – Dist.(m)	LEDs				
	123	124	134	234	1234
LED 1	0,180±0,036	0,160±0,016	0,181±0,033	–	0,158±0,020
LED 2	0,153±0,019	0,154±0,021	–	0,160±0,024	0,154±0,020
LED 3	0,146±0,022	–	0,147±0,021	0,165±0,017	0,145±0,020
LED 4	–	0,135±0,014	0,132±0,012	0,132±0,013	0,132±0,012

### 5.2.2. Regresor SVR Lineal

Después de utilizar el regresor lineal, para poder comparar el funcionamiento y la precisión se decidió usar el regresor SVR lineal. Al igual que en el caso anterior se han realizado con este tipo de regresores dos experimentos: uno para estimar las coordenadas, para lo que se usaban dos regresores uno para la coordenada  $X$  y otro para la coordenada  $Y$ . El otro estimando la distancia de una posición a los LEDs, en este caso se ha usado un regresor por LED, por lo que hay cuatro regresores. Además, al igual que en el regresor lineal, después de esta estimación se ha calculado las coordenadas para las combinaciones de 3 o más LEDs. Según se iba realizando el proceso se ha ido representando los datos en distintas gráficas que se comentaran a continuación y se han incluido los resultados en unas tablas. En esta sección, se encuentran lo más relevante el resto de gráficas y tablas están en la sección de Anexo A.2.

En este caso, también se ha representado las posiciones reales y las estimadas, tanto para el caso de las coordenadas como para las distancias, estas gráficas se pueden ver en las Figuras 5.7a y 5.7b. En este caso, se ve que ambas estimaciones no son tan precisas como el regresor anterior ya que hay una gran variación entre las posiciones reales y las estimadas. En los dos tipos de experimentos, lo que se observa es que los regresores tienden a mover la posición al centro del plano, situando la gran mayoría de las posiciones entre 1,5 y 2m para la coordenada  $X$  y 1 y 2m para la coordenada  $Y$ .



(a) Distribución de los datos reales y los estimados sobre el plano con el experimento tipo 1. (b) Distribución de los datos reales y los estimados sobre el plano con el experimento tipo 2.

Figura 5.7: Distribución de los datos reales y los estimados sobre el plano para el regresor SVR lineal.

En cuanto a la representación de la relación distancia-intensidad para el LED 1, que se puede encontrar en las Figuras 5.8 y 5.9, como en el regresor anterior, la forma de la curva es una parábola y en las gráficas de la estimación se observa que los puntos tienen una mayor dispersión y, por tanto, su curva no es tan precisa. Pero a diferencia del caso anterior observado en las Figuras 5.2 y 5.3, donde había poca diferencia entre la real y la

estimada, en este caso cambia la pendiente de la curva y el rango de la misma, acortándose y aumentando su pendiente. Esto es debido a que las posiciones se ha estimado que están dentro de un área concreta mencionada anteriormente, es decir, se han desplazado al centro y, por tanto, la distancia a los LEDs es muy similar.

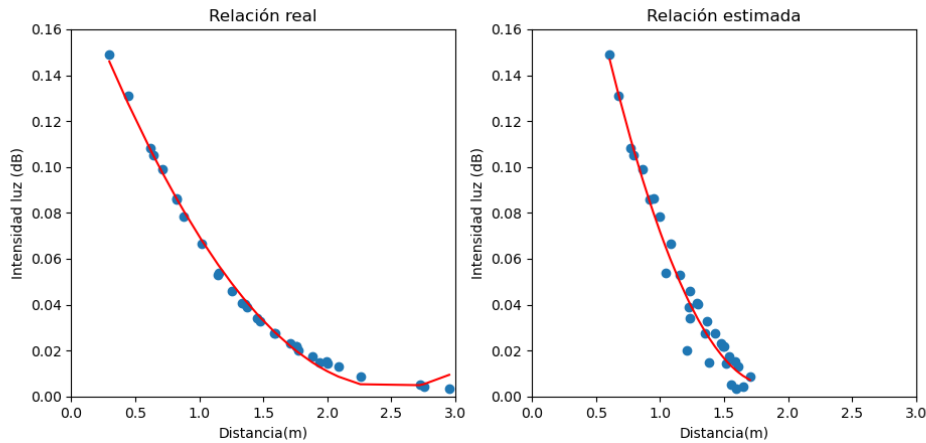


Figura 5.8: Relación Intensidad-Distancia de los datos reales y los estimados para el LED 1 con el experimento tipo 1.

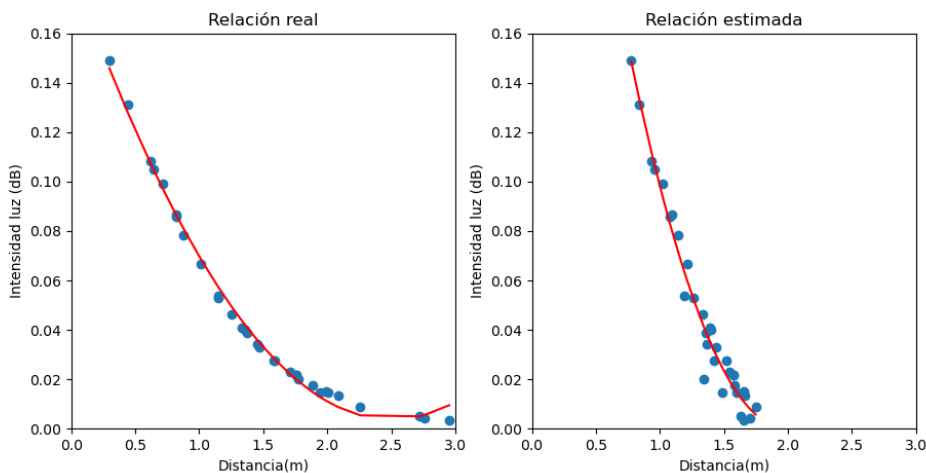


Figura 5.9: Relación Intensidad-Distancia de los datos reales y los estimados para el LED 1 con el experimento tipo 2.

Para el análisis de los datos también se ha llevado a cabo la representación de la CDF del RMSE, que se puede ver en la Figura 5.10. Además, se ha representado la CDF de las distancias a cada LED, que está disponible en la Figura 5.12a. En ambas gráficas, también se ha representado mediante líneas punteadas el percentil 50 y el 90.

Como se puede ver en ambas, su rango es mayor que en el caso de regresor lineal, esto concuerda con la representación de las posiciones sobre el mapa en el que se veía un



gran desplazamiento llegando casi hasta los 0,6m. En el caso de las coordenadas, el menor error está en la coordenada  $Y$  calculada de la distancia y la que peor estimación muestra es la coordenada  $X$  en la estimación de las coordenadas. Asimismo, las coordenadas  $Y$  muestran una estimación más correcta que las coordenadas  $X$ .

En cuanto a los percentiles, para el percentil 50 está entre valores de 0,35m y 0,45m, para el percentil 90, se encuentran entre valores de 0,4m y 0,5m. Para la representación de las distancias a los LEDs, se observa que los LEDs 3 y 4 obtienen relativamente los mismos resultados siendo los mejores y el LED 1 el que peor desempeño realiza con diferencia. En este caso, los percentiles 50, están entorno a valores de 0,35 y 0,4 pero para el percentil 90 el rango aumenta a causa del LED 1 moviéndose entre 0,4m a 0,5m.

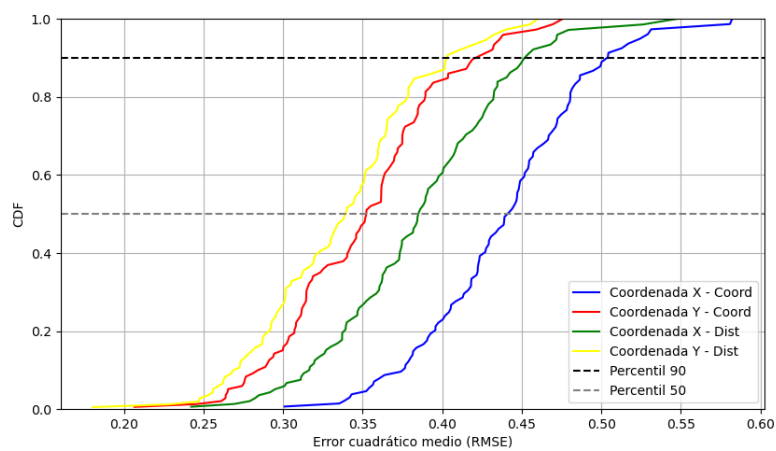


Figura 5.10: CDF del RMSE de las coordenadas ( $x,y$ ) obtenidas para el experimento tipo 1 (coord.) y el experimento tipo 2 (dist.).

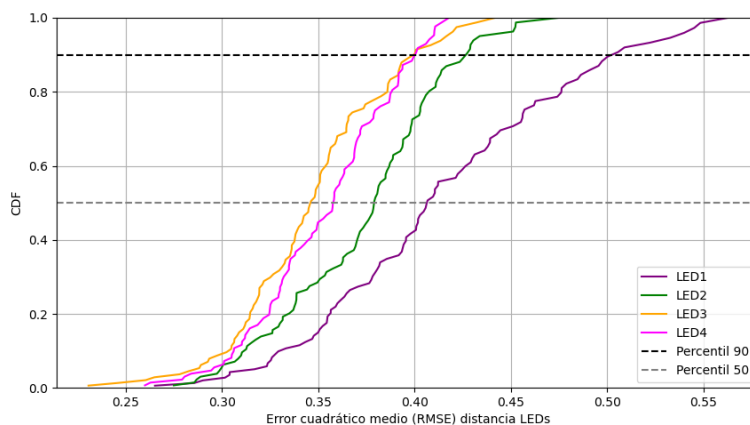
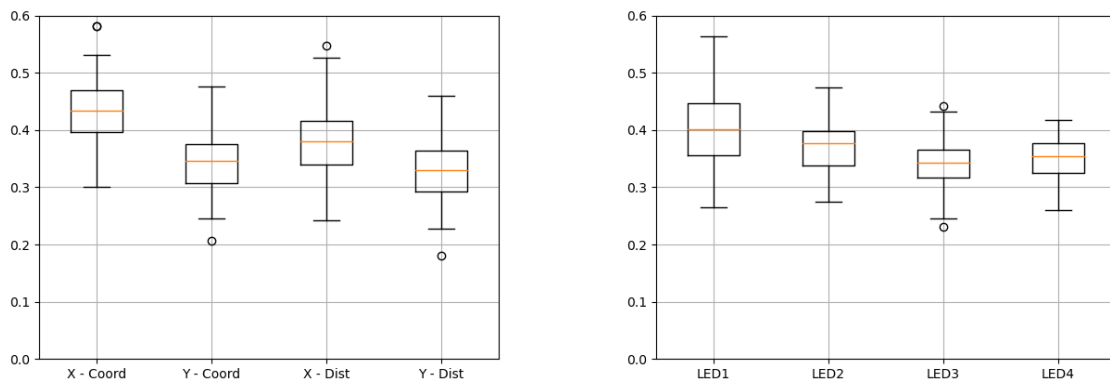


Figura 5.11: CDF del RMSE de las distancias obtenidas para el experimento tipo 2.

Finalizando con las representaciones para este regresor están los diagramas de cajas, que se ven en las Figuras 5.12a y 5.12b. Como con todas las representaciones, se han

separado en dos gráficas según las estimaciones que se han realizado. En el caso de las coordenadas, lo más relevante, a parte de lo comentado en la anterior gráfica y que en esta se muestra también perfectamente, es que todas las coordenadas con sus distintos errores tienen el mismo rango, es decir, el ancho de las cajas es el mismo. En el caso de las distancias, resaltar que el rango del LED 1 es el mayor y el del LED 3 y 4 el menor, confirmaciones que ya se veían en las gráficas de las CDF. Esto sucede debido a que las posiciones a estimar están más alejadas del LED 1 y más cercanas a los LEDs 3 y 4. De tal manera que al realizar la estimación, los valores de los errores de las posiciones más alejadas son mayores que los de las posiciones más cercanas, ya que la información que ofrece la base de datos no es suficiente para que los regresores puedan hacer una estimación igual de precisa para todos los LEDs.



(a) Box-plot del RMSE para el experimento tipo 1. (b) Box-plot del RMSE para el experimento tipo 2.

Figura 5.12: Boxplot del RMSE para el regresor SVR lineal.

Por último, se han recogido en tres tablas los resultados de las distintas estimaciones y cálculos con las mismas combinaciones que el regresor anterior. De tal forma, que el experimento tipo 2 vuelve a obtener mejores resultados que el tipo 1. Además, en el experimento tipo 2 se obtienen menores errores al final del experimento que con la estimación únicamente de las distancias.

Por otro lado, se vuelve a confirmar lo que ya se veía en las otras gráficas en cuanto a qué experimento tenía un menor error y para que LEDs la estimación había sido más acertada. Todas estas tablas se pueden encontrar en las Tablas 5.4, 5.5 y 5.6. Obteniendo el menor error entre 0.48m y 0.36m para la coordenada  $x$  y entre 0.43m y 0.33m para la coordenada  $y$ , ambos en el caso del experimento tipo 2. Además, esto también se observa en los errores MAE y MAER con los menores valores entorno a 0.3m y 0.25m, si se ven las Tablas A.33 y A.36 .

Tabla 5.4: Error cuadrático medio para el regresor svr lineal en el caso del experimento tipo 1 para cada combinación de LEDs.

RMSE – Coord.	LEDs				
	123	124	134	234	1234
Coord. X (m)	0,460±0,047	0,463±0,046	0,447±0,046	0,491±0,049	0,423±0,046
Coord. Y (m)	0,397±0,054	0,425±0,062	0,391±0,059	0,475±0,060	0,348±0,059

Tabla 5.5: Error cuadrático medio para el regresor svr lineal en el caso del experimento tipo 2 para cada combinación de LEDs.

RMSE – Dist.(m) (Coord.)	LEDs				
	123	124	134	234	1234
Coord. X (m)	0,481±0,054	0,399±0,050	0,443±0,046	0,469±0,060	0,366±0,047
Coord. Y (m)	0,419±0,049	0,395±0,057	0,406±0,067	0,434±0,068	0,335±0,059

Tabla 5.6: Error cuadrático medio para el regresor svr lineal en el caso de estimar las distancias para cada combinación de LEDs.

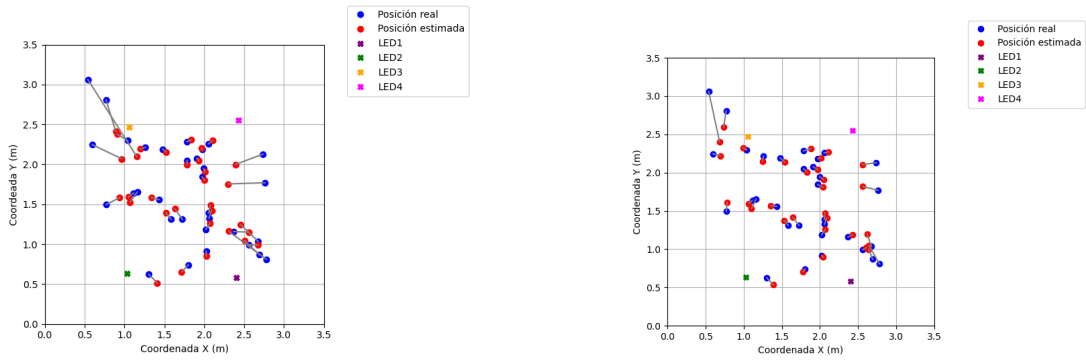
RMSE – Dist.(m)	LEDs				
	123	124	134	234	1234
LED 1	0,404±0,057	0,483±0,063	0,402±0,062	–	0,402±0,060
LED 2	0,410±0,040	0,382±0,042	–	0,374±0,041	0,372±0,041
LED 3	0,353±0,043	–	0,355±0,041	0,477±0,051	0,353±0,042
LED 4	–	0,341±0,035	0,371±0,036	0,354±0,036	0,342±0,035

### 5.2.3. Regresor SVR polinomial

Tras los anteriores regresores se ha procedido a repetir el mismo proceso con el regresor RBF polinomial. A continuación, se ven las distintas gráficas y algunos de los resultados que se han obtenido del experimento.

Para comenzar, la primera representación que se ha realizado es la de la estimación de las posiciones para ambos tipos de experimentos. Como se puede ver en las Figuras 5.13a y 5.13b, principalmente en las posiciones centrales del plano la estimación, correspondiente con los puntos rojos, coincide casi a la perfección con la posición real, puntos azules. A medida que las posiciones se desplazan hacia los laterales del plano, la estimación empieza a no ser tan precisa y el desempeño de los regresores se ve afectado.

Por otro lado, si se comparan los dos tipos de experimentos entre sí, se puede apreciar que el experimento tipo 2 ofrece mejores resultados que los de las coordenadas, sobre todo para las posiciones que están situadas en el centro del plano. Esto como se ha demostrado anteriormente ocurre también en el resto de regresores.



(a) Distribución de los datos reales y los estimados sobre el plano con el experimento tipo 1. (b) Distribución de los datos reales y los estimados sobre el plano con el experimento tipo 2.

Figura 5.13: Distribución de los datos reales y los estimados sobre el plano para el regresor SVR polinomial.

La siguiente representación se trata de la relación distancia-intensidad como se puede ver en las Figuras 5.14 y 5.15. Como todas las anteriores la forma de la curva de aproximación es una parábola cóncava. Comparando las gráficas de las reales con las estimaciones vemos que la forma es la misma, así como la pendiente de todas ellas.

Para este regresor se observa lo que ya se vio con el regresor lineal y es que las distancias de mayor longitud desaparecen, por lo que la curva de aproximación acaba antes. Esto coincide con el hecho de que las posiciones más alejadas del centro en las estimaciones, han sido desplazadas al centro, reduciendo así la distancia a los LEDs. Además, cabe destacar que en este caso las dos estimaciones presentan una gran dispersión en cuanto a los valores de relación y la curva de aproximación, hecho que no sucedía o en mucha menor medida para los anteriores regresores. Aun así, como se ha dicho la curva de aproximación en la estimación se corresponde casi con exactitud con la de las posiciones reales.

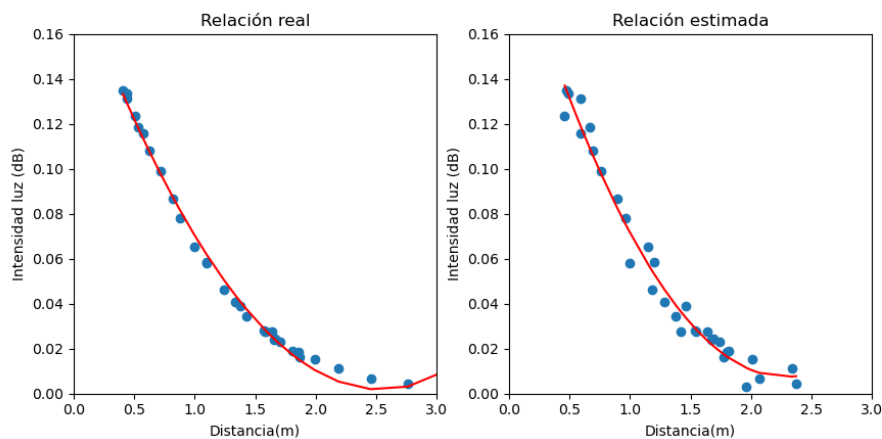


Figura 5.14: Relación Intensidad-Distancia de los datos reales y los estimados para el LED 1 con el experimento tipo 2.

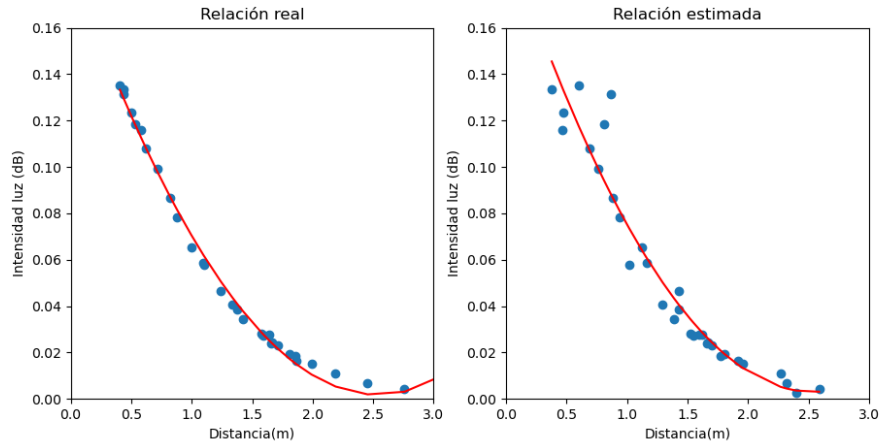


Figura 5.15: Relación Intensidad-Distancia de los datos reales y los estimados para el LED 1 con el experimento tipo 2.

Otra de las gráficas que se han representado es la de la CDF del RMSE. Como en los casos anteriores se tienen por un lado las coordenadas y, por otro, las distancias a los LEDs. Todas estas gráficas se pueden encontrar en las Figuras 5.16 y 5.17.

En el caso de las coordenadas, se observa que los resultados son mejores que el caso anterior, ya que los valores del RMSE son menores. Así, se puede ver que las coordenadas calculadas mediante el experimento tipo 2 son más precisas a la hora de obtener la posición que las del experimento tipo 1. Además, con los percentiles se confirma todo lo dicho, para el percentil 50 los valores son entre 0.1m y 0.2m y para el percentil 90 entre 0.15m y 0.25m. En el caso de las distancias, se observa que los resultados son aproximadamente iguales que los de las coordenadas calculadas en términos de los valores del RMSE. Siendo el LED con mejores resultados el 4 y el 1 el de peores resultados. También analizando los percentiles se confirma este funcionamiento más preciso, con valores entre 0.1m y 0.18m para el percentil 50 y para el percentil 90 entre 0.12m y 0.22m.

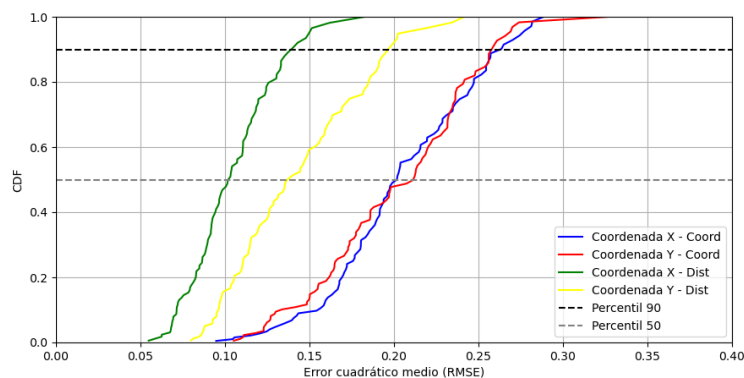


Figura 5.16: CDF del RMSE de las coordenadas  $(x,y)$  obtenidas para el experimento tipo 1 (coord.) y el experimento tipo 2 (dist.).

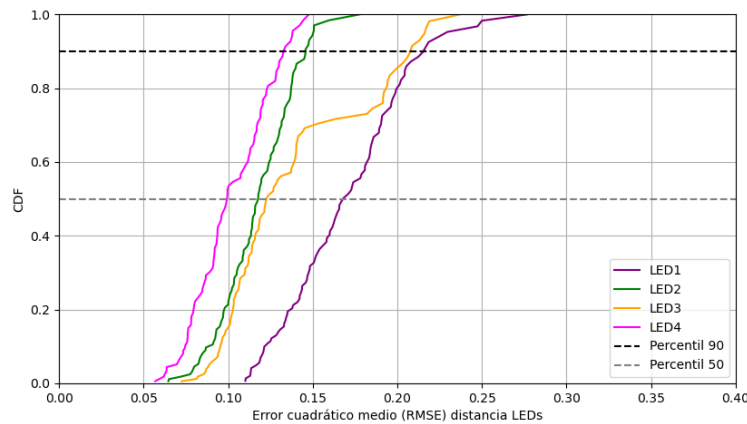
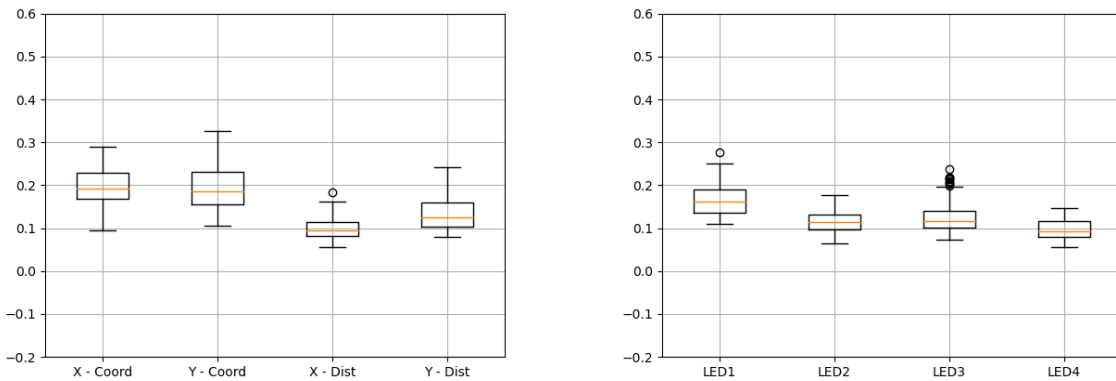


Figura 5.17: CDF del RMSE de las distancias obtenidas para el experimento tipo 2.

Finalmente, las últimas representaciones que se encuentran son los diagramas de cajas disponibles en las Figuras 5.18a y 5.18b. Como se puede observar, confirman todo lo que se ha mencionado anteriormente de que LED funciona mejor, así como en la comparación entre experimentos. Lo más interesante a mencionar es que su desempeño se parece mucho al del regresor lineal como se puede ver al comparar los diagramas. A parte de que para el LED 3 aparecen varios valores atípicos de RMSE.



(a) Box-plot del RMSE para el experimento tipo 1.

(b) Box-plot del RMSE para el experimento tipo 2.

Figura 5.18: Boxplot del RMSE para el regresor SVR polinomial.

Por último, se han incluido los resultados más importantes en las Tablas 5.7, 5.8 y 5.9. De tal forma que con este regresor aparecen algunos de los menores valores de RMSE vistos hasta ahora. Dentro de las coordenadas la estimada con mayor precisión es la que proviene del experimento tipo 2 siendo la coordenada  $X$  con un valor de 0.1m. Obteniéndose unos valores mayores de RMSE para el experimento tipo 1 con valores entorno a los 0.2m. Además, dentro de las distancias se repiten las conclusiones anteriores, el LED 4 es el que

mejor resultados obtiene y el LED 1 el que refleja una estimación menos precisa que el resto y la combinación de los LEDs 1-2-3-4 nos aporta la mejor estimación de la posición del receptor. Viendo las Tablas A.51 y A.54 que se encuentran en el anexo también se observa esta tendencia con valores entorno a 0.13m y 0.1m para los errores de MAE y MAER.

Tabla 5.7: Error cuadrático medio para el regresor svr polinomial en el caso del experimento tipo 1 para cada combinación de LEDs.

RMSE – Coord.	LEDs				
	123	124	134	234	1234
Coord. X (m)	0,307±0,082	0,249±0,049	0,235±0,054	0,323±0,087	0,200±0,053
Coord. Y (m)	0,229±0,048	0,244±0,052	0,284±0,073	0,322±0,088	0,195±0,051

Tabla 5.8: Error cuadrático medio para el regresor svr polinomial en el caso del experimento tipo 2 para cada combinación de LEDs.

RMSE – Dist.(m) (Coord.)	LEDs				
	123	124	134	234	1234
Coord. X (m)	0,306±0,104	0,203±0,073	0,221±0,067	0,195±0,093	0,100±0,032
Coord. Y (m)	0,223±0,053	0,214±0,065	0,282±0,090	0,212±0,063	0,135±0,041

Tabla 5.9: Error cuadrático medio para el regresor svr polinomial en el caso de estimar las distancias para cada combinación de LEDs.

RMSE – Dist.(m)	LEDs				
	123	124	134	234	1234
LED 1	0,229±0,051	0,192±0,040	0,243±0,057	–	0,169±0,039
LED 2	0,125±0,022	0,172±0,033	–	0,171±0,033	0,109±0,023
LED 3	0,176±0,04	–	0,177±0,036	0,212±0,045	0,137±0,043
LED 4	–	0,165±0,029	0,119±0,027	0,154±0,028	0,095±0,022

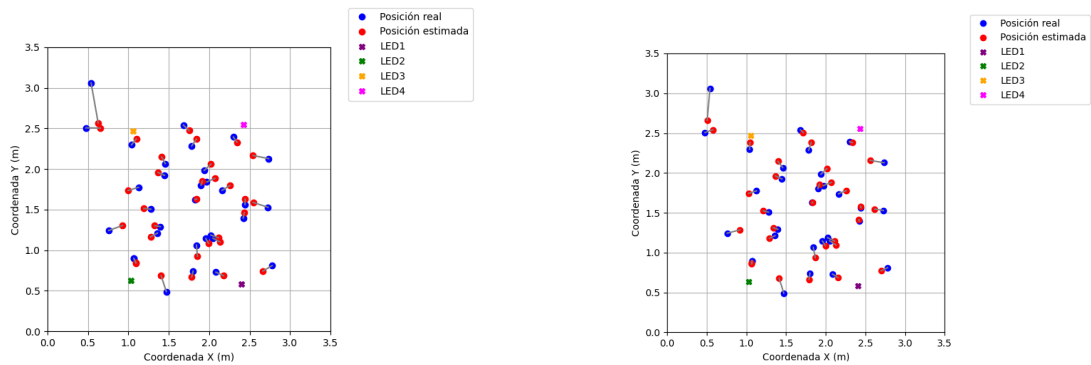
#### 5.2.4. Regresor SVR RBF

Finalmente, se ha realizado los últimos tipos de experimentos con el regresor SVR RBF. Al igual que en los anteriores regresores, tras los experimentos se han recogido los datos y se han representado tanto con gráficas como con tablas.

En primer lugar, se encuentra la representación de las posiciones reales en comparación con las estimadas. Ambas gráficas se pueden ver en las Figuras 5.19a y 5.19b. Con este regresor, se puede observar que su desempeño es muy preciso, ya que apenas hay diferencia entre los valores reales, puntos azules, y los valores estimados, puntos rojos. Más adelante, con los resultados numéricos se podrá aclarar, pero este regresor es el más preciso de todos

los usados a la hora de la estimación de la posición. Además, a diferencia del anterior regresor, el regresor SVR polinomial, en el que las posiciones que no estaban situadas en el centro del plano variaban bastante sus posiciones. Excepto por algunos valores concretos, en general hay muy poca diferencia entre los valores reales y los estimados.

También, si se comparan los dos tipos de experimentos, se puede observar algo que ya se ha comprobado con el resto de regresores, y es que el experimento tipo 2 es más preciso que el experimento tipo 1.



(a) Distribución de los datos reales y los estimados sobre el plano con el experimento tipo 1. (b) Distribución de los datos reales y los estimados sobre el plano con el experimento tipo 2.

Figura 5.19: Distribución de los datos reales y los estimados sobre el plano para el regresor SVR RBF.

Pasando a las siguientes gráficas, éstas son la relación distancia-intensidad, que se pueden ver en las Figuras 5.20 y 5.21. Como todos los anteriores regresores, la curva de aproximación forma una parábola cóncava. Al comparar las gráficas de los datos reales con las estimaciones, se observa para ambos tipos de experimentos unas gráficas muy parecidas. Con la misma pendiente y siendo la mayor diferencia el rango de las parábolas debido a la estimación de las posiciones más alejadas del centro. Aún así, el experimento tipo 2 obtiene un rango mayor que la de las coordenadas, por lo que es más precisa. Otro dato a destacar es que los valores de las estimaciones coinciden en gran medida con la curva de aproximación, es decir, no hay tanta dispersión como en los regresores anteriores.



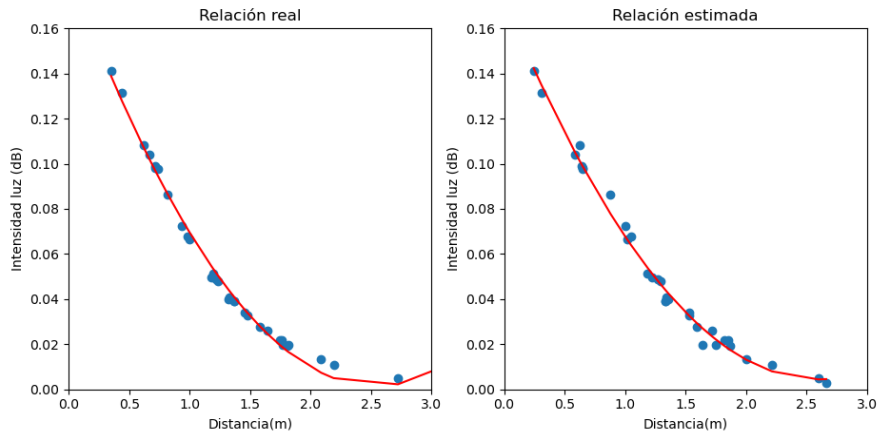


Figura 5.20: Relación Intensidad-Distancia de los datos reales y los estimados para el LED 1 con el experimento tipo 1.

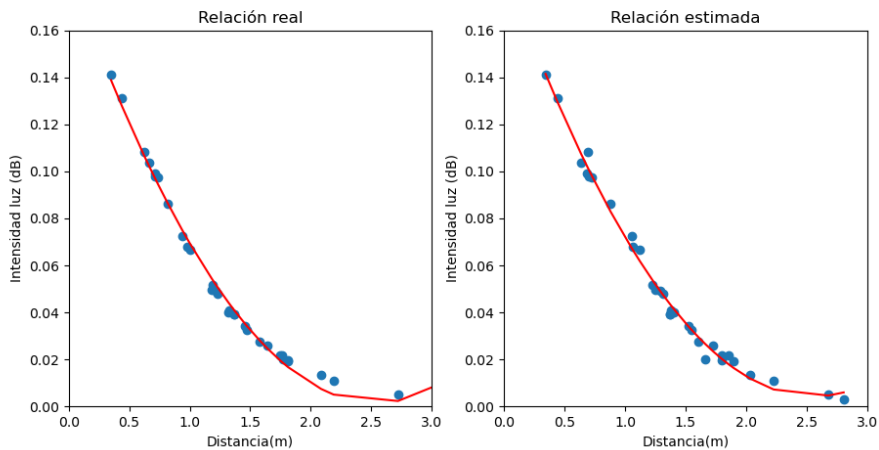


Figura 5.21: Relación Intensidad-Distancia de los datos reales y los estimados para el LED 1 con el experimento tipo 2.

A continuación, se han calculado las CDF del RMSE de las distintas estimaciones y es lo que se presenta en las Figuras 5.22 y 5.23. Por un lado, las coordenadas  $y$ , y por otro, las distancias a los LEDs. En el caso de las coordenadas, un dato a remarcar es que todas las CDF empiezan desde el mismo valor aproximadamente y a medida que ésta va aumentando se dispersan más sus valores, siendo la coordenada  $X$  obtenida del experimento tipo 2 la que menor RMSE tiene y la coordenada  $Y$  obtenida del experimento tipo 1, la que peor estimación realiza. En cuanto a los percentiles se mueven entre valores de 0.08m y 0.12m para el percentil 50 y 0.1m y 0.18m para el percentil 90, lo que demuestra que este regresor es el que mayor precisión ha demostrado tener a la hora de realizar las estimaciones.

Por el lado de las distancias, todos los LEDs han realizado unas estimaciones igual de precisas entre si y con un desempeño igual al tipo de experimento de la estimación de las

coordenadas. Esto lo vemos en sus valores de percentil, teniendo para el percentil 50 un valor entorno a 0.1m y para el percentil 90 entre 0.1m y 0.15m.

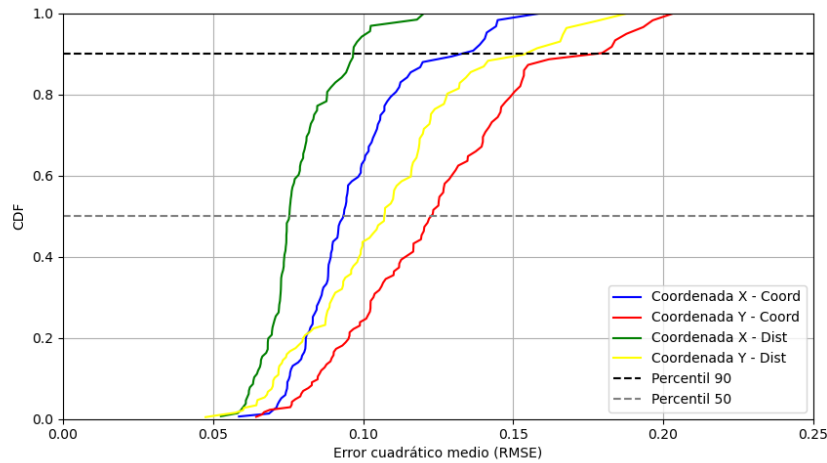


Figura 5.22: CDF del RMSE de las coordenadas ( $x,y$ ) obtenidas para el experimento tipo 1 (coord.) y el experimento tipo 2 (dist.).

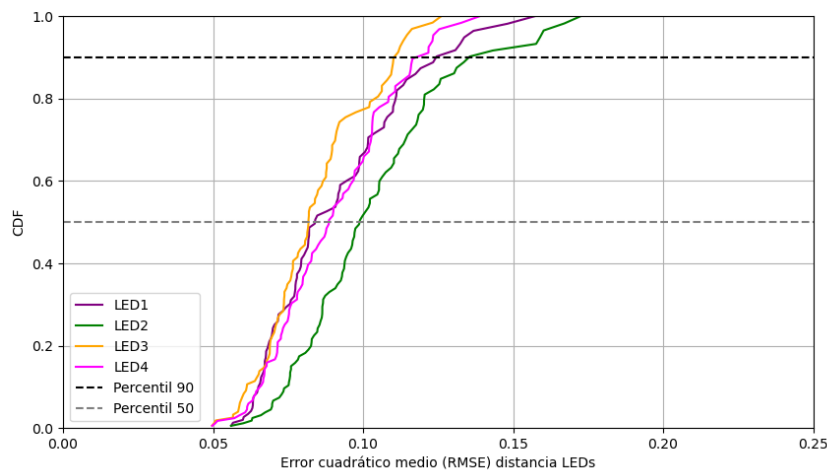
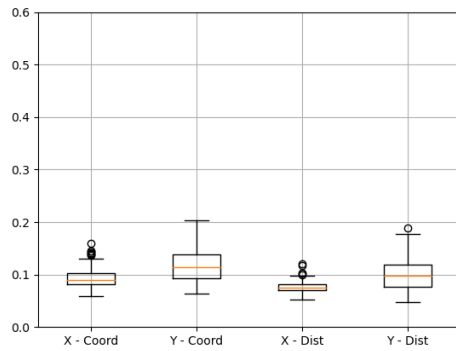
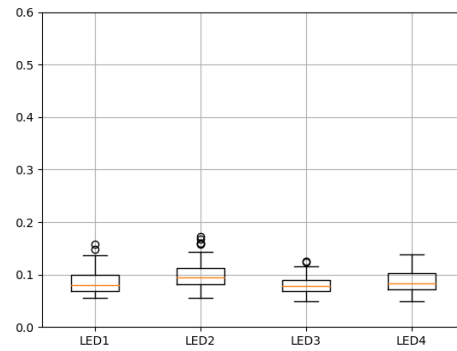


Figura 5.23: CDF del RMSE de las distancias obtenidas para el experimento tipo 2.

Finalmente, se encuentran las gráficas que representan los diagramas de cajas como se puede observar en las Figuras 5.24a y 5.24b. Como se puede ver en la comparación entre coordenadas y distancias, ambos experimentos obtienen resultados similares. En el caso de las coordenadas, destacar que la coordenada  $X$  calculada de las dos formas distintas tiene menor rango que el resto. Además, la estimación de las distancias ha dado unos resultados más precisos en comparación con los regresores anteriores. Aunque esto se puede observar en mayor o menor medida en todos las cajas.



(a) Box-plot del RMSE para el experimento tipo 1.



(b) Box-plot del RMSE para el experimento tipo 2.

Figura 5.24: Box-plot de los RMSE para el regresor SVR RBF.

Para acabar, se han recogido los datos del RMSE y se han concentrado en tres tablas que están representadas en las Tablas 5.10, 5.11 y 5.12. Al ver los números, lo primero a destacar es lo que se ha ido mostrando a lo largo de los experimentos con este regresor y es que es el que mejores estimaciones ha realizado, ya que el RMSE en todos los casos es el menor, todos con valores entorno a 0.1m algunos incluso por debajo. También mencionar que para este regresor las coordenadas  $X$  han sido las más precisas, con valores entre los 0.07m y 0.16m.

En cuanto a las distancias, todas se encuentran en los mismos valores aproximadamente de RMSE, excepto el LED 2 que tiene un error algo mayor que el resto, pero no hay entre ellos una gran diferencia, encontrándose todos estos valores entorno a los 0.1m.

En cuanto a las otras métricas, los menores de valores de MAE se encuentran entorno a 0.07m y para el MAER con valores cerca de los 0.06m. Todos ellos para el experimento tipo 2 que es el que ofrece una estimación más precisa. Se pueden ver las tablas A.69 y A.72 que se encuentran en el Anexo.

Tabla 5.10: Error cuadrático medio para el regresor SVR rbf en el caso del experimento tipo 1 para cada combinación de LEDs.

RMSE – Coord.	LEDs				
	123	124	134	234	1234
Coord. X (m)	$0,173 \pm 0,060$	$0,103 \pm 0,021$	$0,127 \pm 0,030$	$0,112 \pm 0,038$	$0,092 \pm 0,018$
Coord. Y (m)	$0,119 \pm 0,026$	$0,117 \pm 0,029$	$0,152 \pm 0,047$	$0,123 \pm 0,035$	$0,121 \pm 0,03$

Tabla 5.11: Error cuadrático medio para el regresor SVR rbf en el caso del experimento 2 para cada combinación de LEDs.

RMSE – Dist.(m) (Coord.)	LEDs				
	123	124	134	234	1234
Coord. X (m)	0,16±0,052	0,093±0,019	0,119±0,034	0,087±0,016	0,075±0,01
Coord. Y (m)	0,109±0,025	0,106±0,026	0,15±0,051	0,107±0,036	0,103±0,028

Tabla 5.12: Error cuadrático medio para el regresor SVR rbf en el caso de estimar las distancias para cada combinación de LEDs.

RMSE – Dist.(m)	LEDs				
	123	124	134	234	1234
LED 1	0,105±0,028	0,085±0,02	0,125±0,036	–	0,089±0,02
LED 2	0,086±0,016	0,1±0,023	–	0,101±0,029	0,1±0,023
LED 3	0,091±0,022	–	0,093±0,02	0,087±0,013	0,082±0,02
LED 4	–	0,085±0,019	0,076±0,01	0,079±0,011	0,085±0,02

### 5.3. Comparativa

En esta sección, para concluir con este apartado se han recogido todos los valores de RMSE en tres tablas según los experimentos que se han realizado cada una con los distintos regresores para tener una comparación. Estas tablas solo muestran los datos para la combinación de LEDs 1-2-3-4 y se pueden ver los datos en las Tablas 5.13, 5.14 y 5.15.

De esta forma, se puede apreciar qué regresores ofrecen un mejor rendimiento. Así, para el experimento tipo 1 donde directamente se estiman las coordenadas del receptor, ordenándolos según como sea su rendimiento, el mejor regresor se corresponde con el regresor SVR RBF, seguido por los regresores lineal y el regresor SVR polinomial, que ofrecen unos resultados muy similares y, por último, el regresor SVR lineal. Siendo el RMSE, cuatro veces menor en el regresor SVR RBF comparándolo con el regresor SVR lineal. Esto sucede ya que el regresor SVR RBF puede encontrar de una forma más precisa las relaciones entre los datos a diferencia del resto y, así, adaptarse a las distintas variaciones que sufren los datos permitiendo ofrecer mejores resultados.

Además, comparando entre los dos tipos de experimentos para la obtención de las coordenadas, se observa en la tablas 5.13 y 5.14 lo que se ha comentado a lo largo de toda la sección y es que las coordenadas obtenidas en el experimento tipo 2 son más precisas que las coordenadas obtenidas en el tipo 1. En este caso, el rendimiento de los regresores es igual que el caso anterior en términos de cual es mejor.

Por último, para la estimación de las distancias, en general hay algunas distinciones entre los LEDs como vemos en la Tabla 5.15. Por ejemplo, los LEDs 3 y 4, principalmente

este último, son los que obtienen mejores resultados de RMSE y los LEDs 1 y 2, particularmente el LED 1, los que obtienen unos valores más altos de RMSE. Cabe mencionar que estas afirmaciones no son tan evidentes como en el caso de las coordenadas. Ya que en este caso algunos regresores ofrecen mejores resultados para unos LEDs y peores para otros. Esto se debe a la cantidad de posiciones que haya cerca de un LED y a que distancia están esas posiciones.

Tabla 5.13: Errores cuadráticos medios para todos los regresores en el caso del experimento tipo 1 para la combinación de los 4 LEDs.

RMSE – Dist.(m)	Tipo de regresor			
	Reg. lineal	Reg. SVR lineal	Reg. SVR polinomial	Reg. SVR rbf
Coord. X (m)	0,177±0,029	0,423±0,046	0,200±0,053	0,092±0,018
Coord. Y (m)	0,15±0,032	0,348±0,059	0,195±0,051	0,121±0,03

Tabla 5.14: Errores cuadráticos medios para todos los regresores en el caso del experimento tipo 2 para la combinación de los 4 LEDs.

RMSE – Dist. (m) (coord.)	Tipos de regresores			
	Reg. lineal	Reg. SVR lineal	Reg. SVR polinomial	Reg. SVR rbf
Coord. X (m)	0,108±0,014	0,366±0,047	0,100±0,032	0,075±0,01
Coord. Y (m)	0,155±0,024	0,335±0,059	0,135±0,041	0,103±0,028

Tabla 5.15: Errores cuadráticos medios para todos los regresores en el caso de estimar las distancias para la combinación de los 4 LEDs.

RMSE – Dist.(m)	Tipo de regresor			
	Regresor lineal	Regresor SVR lineal	Regresor SVR polinomial	Regresor SVR rbf
LED 1	0,158±0,020	0,402±0,060	0,169±0,039	0,089±0,02
LED 2	0,154±0,020	0,372±0,041	0,109±0,023	0,1±0,023
LED 3	0,145±0,020	0,353±0,042	0,137±0,043	0,082±0,02
LED 4	0,132±0,012	0,342±0,035	0,095±0,022	0,085±0,02

# Capítulo 6

## Conclusiones y Líneas futuras

En esta sección, se comentan las principales conclusiones obtenidas tras la realización de este TFG. También se explica el impacto que tiene analizándolo desde distintos niveles. Además, se detallan las limitaciones encontradas durante la realización del mismo. Así como, posibles líneas futuras a seguir si se quisiera continuar con su desarrollo.

### 6.1. Conclusiones

Durante la realización de este TFG, se ha estudiado cómo estimar las coordenadas del receptor en distintas posiciones utilizando la intensidad de la luz recibida de cuatro LEDs distintos mediante el uso de ML y un receptor con un fotodiodo colocado en el objeto de interés. Además, en la realización se buscaba obtener de entre una serie de regresores cual de ellos desempeñaba un mejor rendimiento a la hora de estimar las coordenadas.

Durante la realización de los experimentos, se ha podido observar que la calidad de los datos utilizados para el entrenamiento de los algoritmos de ML afecta a la salida. Como se ha comentado anteriormente, existían una serie de datos que dada su casuística al utilizarse en el grupo de test, los resultados se verían perjudicados. Así, aunque los datos usados para cada grupo se han seleccionado aleatoriamente, cabe destacar la necesidad de disponer de unos datos adecuados. También destacar que, para evitar que la respuesta dependiera de los datos seleccionados, se han realizado los experimentos un total de 100 veces para obtener una evaluación estadísticamente significativa. De la misma forma, se quería comprobar la importancia de cada LED y como la disponibilidad de una mayor cantidad de datos puede afectar a los resultados. Para ello, se han analizado distintas combinaciones de LEDs. Quedando claro tras los resultados que, a mayor número de datos de entrada, mejores son los resultados. Obteniendo el mejor rendimiento para la combinación de los cuatro LEDs. Por ejemplo, en el caso del experimento tipo 2 para el

regresor SVR se observó un error de 0,089m o 0,082m en el LED 1 y 3 respectivamente con 4 LEDs frente a 0,105m o 0,091m en el caso de 3 LEDs, ver la Tabla 5.12.

Por otro lado, se ha buscado mediante dos tipos de experimentos la forma más precisa de obtener las coordenadas. Viendo los resultados, el experimento tipo 2 junto con el uso del algoritmo de LS para la intersección de circunferencias y, así, calcular las coordenadas del receptor es la mejor opción. Por ejemplo, los resultados para el regresor SVR RBF, en el caso del experimento tipo 1 son errores de 0,092m y 0,121m para las coordenadas  $x$  e  $y$  respectivamente. Sin embargo, para esas coordenadas en el experimento tipo 2 se han obtenido errores de 0,075m y 0,103m, ver las Tablas 5.11 y 5.10.

Finalmente, con la aplicación de distintos regresores a los experimentos se ha respondido a la cuestión de qué regresor desempeñaba mejor el papel de obtener las coordenadas en una posición, siendo este el regresor SVR RBF. Entre los resultados del RMSE de mayor valor y el de menor se encuentra una diferencia de hasta casi 5 veces menor entre uno y otro. Esto se ve en la Tabla 5.14 en la que para el experimento tipo 2 el error del regresor SVR lineal para la coordenada  $x$  es 0.366m, pero para esa coordenada el regresor SVR RBF obtiene un error de 0,075m. Todos los resultados obtenidos a lo largo de este TFG se encuentran en las Tablas comparativas en 5.13, 5.11 y 5.10.

Por último, tras este análisis se puede afirmar que se puede obtener de forma precisa las coordenadas de una posición mediante la intensidad de la luz recibida. Así como, la efectividad del proceso llevado a cabo y de los regresores usados para ello.

## 6.2. Impacto

En esta sección, se explica cual es el impacto que tiene este TFG a nivel social, medioambiental y económico.

### 6.2.1. Impacto social

Gracias a que se ha demostrado que la estimación de las coordenadas de un fotorreceptor en una posición es posible, esta tecnología se puede aplicar en multitud de aplicaciones para saber la ubicación de un objeto o persona en interiores, lo que tiene un gran impacto social. De esta forma, la vida de las personas mejoraría y la haría más cómoda y eficiente. Permitiendo obtener unos resultados muy precisos de una forma sencilla.

Algunos ejemplos de aplicación inmediata son ayudar a personas con algún problema de visión que, gracias a este sistema, sepan donde se encuentran dentro de un edificio, como hospitales o estaciones, y ofrecerles una ruta para llegar al lugar deseado. Esta tecnología también es atractiva para empresas de logística en la ubicación de paquetes o

de máquinas que, sin conductor, puedan ubicarse y moverse dentro de una fábrica, por ejemplo, agilizando así ciertos procesos. Otro caso de uso sería para mejorar la experiencia de un cliente en un comercio pudiendo ofreciéndole la mejor ruta a seguir para encontrar un producto concreto que esté buscando. Además, se podría usar en parkings tanto en la búsqueda de espacio libre indicando la posición exacta del lugar como en la situación de que no se recuerde donde está el vehículo propio y se necesite recordar su ubicación. Asimismo, se podría aplicar en el ámbito de la seguridad, por ejemplo, en prisiones para saber la ubicación exacta de los presos y tenerlos controlados. Pero también en entidades públicas se podría aplicar su uso para comprobar ante un ataque donde se encuentran los agentes o los funcionarios y también donde se encuentran los atacantes.

### 6.2.2. Impacto medioambiental

Esta tecnología necesita de muy pocos recursos, por lo que tendría un impacto ambiental positivo, en lo que se refiere a invadir zonas verdes en su despliegue o al consumo energético que pueda provocar su uso. Por un lado, se necesitan luces LEDs para llevar a cabo el proceso, lo que supone un ahorro energético, que se traduce en una menor contaminación por la obtención de la energía. Además, favoreciendo al medioambiente, ya que las empresas o usuarios que quieran utilizar esta energía y no dispongan en sus instalaciones de este tipo de luces LEDs deben actualizarlas.

Por otro lado, al ser una tecnología muy sencilla de hacer funcionar solo se necesita un receptor con un fotodiodo, lo que supone un coste energético mínimo. Cabe mencionar que el proceso de estimación requiere ciertos recursos informáticos para poder realizar las estimaciones, lo que si que tendría cierto impacto en el medioambiente.

### 6.2.3. Impacto económico

En cuanto al impacto económico, se encuentra como se ha mencionado que es una tecnología sencilla, ya que no se requiere un gran desembolso económico para obtenerla. Además, mediante su uso, se puede obtener un ahorro económico ya sea por el uso de luces LEDs que son de bajo consumo como por la eficiencia adquirida con su uso obteniendo así un menor consumo de otros recursos y una mayor agilidad en ciertos procesos.

A destacar que este TFG se enmarca dentro del Objetivos de Desarrollo Sostenible (ODS)), establecidos en 2015 por la ONU. Son un conjunto de 17 objetivos globales adoptados como parte de la Agenda 2030 para el Desarrollo Sostenible. Éstos se implantaron con el fin de abordar los desafíos más urgentes a los que nos enfrentamos como sociedad promoviendo así la igualdad y el cuidado del medioambiente entre otros. En la Tabla 6.1 se encuentran los objetivos que se ven beneficiados con el estudio realizado en este TFG.



Tabla 6.1: Metas específicas de los ODS que se adecuan a este TFG.

<b>Objetivo</b>	<b>Meta</b>	<b>Descripción</b>
ODS 7: Garantizar el acceso a una energía asequible, segura, sostenible y moderna	ODS 7.3	De aquí a 2030, duplicar la tasa mundial de mejora de la eficiencia energética
	ODS 7.b	De aquí a 2030, ampliar la infraestructura y mejorar la tecnología para prestar servicios energéticos modernos y sostenibles para todos en los países en desarrollo, en particular los países menos adelantados, los pequeños Estados insulares en desarrollo y los países en desarrollo sin litoral, en consonancia con sus respectivos programas de apoyo
ODS 8: Promover el crecimiento económico inclusivo y sostenible, el empleo y el trabajo decente para todos	ODS 8.2	Lograr niveles más elevados de productividad económica mediante la diversificación, la modernización tecnológica y la innovación, entre otras cosas centrándose en los sectores con gran valor añadido y un uso intensivo de la mano de obra
ODS 9: Construir infraestructuras resilientes, promover la industrialización sostenible y fomentar la innovación	ODS 9.1	Desarrollar infraestructuras fiables, sostenibles, resilientes y de calidad, incluidas infraestructuras regionales y transfronterizas, para apoyar el desarrollo económico y el bienestar humano, haciendo especial hincapié en el acceso asequible y equitativo para todos
	ODS 9.4	De aquí a 2030, modernizar la infraestructura y reconvertir las industrias para que sean sostenibles, utilizando los recursos con mayor eficacia y promoviendo la adopción de tecnologías y procesos industriales limpios y ambientalmente racionales, y logrando que todos los países tomen medidas de acuerdo con sus capacidades respectivas
ODS 11: Lograr que las ciudades sean más inclusivas, seguras, resilientes y sostenibles	ODS 11.3	De aquí a 2030, aumentar la urbanización inclusiva y sostenible y la capacidad para la planificación y la gestión participativas, integradas y sostenibles de los asentamientos humanos en todos los países
ODS 12: Garantizar modalidades de consumo y producción sostenibles	ODS 12.6	Alentar a las empresas, en especial las grandes empresas y las empresas transnacionales, a que adopten prácticas sostenibles e incorporen información sobre la sostenibilidad en su ciclo de presentación de informes
	ODS 12.a	Ayudar a los países en desarrollo a fortalecer su capacidad científica y tecnológica para avanzar hacia modalidades de consumo y producción más sostenibles

### 6.3. Lecciones aprendidas

Durante la realización de TFG, se han aprendido ciertas destrezas y se han ampliado conocimientos sobre temas conocidos anteriormente. A continuación, se exponen cada uno de ellos:

- Se han adquirido conocimientos sobre distintas técnicas con las que asegurarse de que se obtienen unos resultados válidos como validación cruzada.
- Se han expandido los conocimientos que se tenían sobre el lenguaje de *Python* a través de las distintas clases utilizadas.
- Mediante el uso de algoritmos de ML se han afianzado competencias sobre el funcionamiento de los regresores.
- Se ha aprendido a realizar búsquedas en bases de datos como Google Scholar para obtener información de literatura científica.
- Para la redacción de este TFG se ha aprendido el lenguaje de LaTeX para poder aplicar así las distintas herramientas que éste ofrece.

### 6.4. Líneas futuras

El estudio del posicionamiento mediante luz visible es un campo en el que existen grandes posibilidades de investigación y desarrollo. Continuando con el estudio llevado a cabo durante este TFG se podría explorar la posibilidad de usar otro tipo de regresores diferentes como *Ridge Regression* o *Lasso Regression* a los ya utilizados con el fin de ver su rendimiento y poder tener así una comparación con más resultados.

También se encuentran distintas rutas por las que seguir estudiando en cuanto al espacio usado, como escenarios más complejos con oclusiones parciales, etc. Ya que este tipo de tecnología es utilizada en interiores y para este estudio tenemos un espacio con unas medidas concretas, pudiendo variar las mismas o incluso la posición de los LEDs utilizados.

De la misma forma, los datos para este TFG se han obtenido para el caso en el que se tengan 4 LEDs pudiendo variar este número y, así, explorar diferentes posibilidades observando si al aumentar o disminuir el número de LEDs el rendimiento de los regresores se ve afectado.

Por último, a lo largo del TFG se han explicado distintas técnicas de validación cruzada, pero solo se ha utilizado una, en este caso se ha usado *k-fold*. Así, se podrían emplear

---

cualquiera de las otras técnicas que existen y comparar cual de ellas tiene mayor eficacia entorno a la cuestión de poder obtener información sobre cual regresor es más preciso. Por ejemplo, se podría aplicar *Stratified Cross-Validation* o la técnica *Jackknife*, ver más técnicas de validación cruzada en 3.4.

# Apéndice A

## Apéndice

### A.1. Capítulo 4 - Metodología

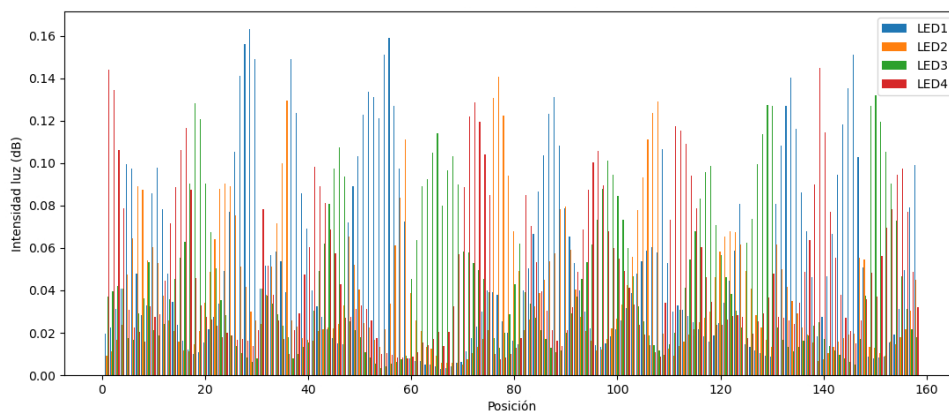


Figura A.1: Histograma Intensidad-Frecuencia por colores para cada LED.

### A.2. Capítulo 5 - Resultados

#### A.2.1. Tablas del Regresor lineal

Tabla A.1: Error cuadrático medio para el regresor lineal en el caso del experimento tipo 1 para cada combinación de LEDs 1.

RMSE – Coord.	LEDs			
	1	2	3	4
Coord. X (m)	0,445±0,046	0,549±0,059	0,453±0,048	0,543±0,054
Coord. Y (m)	0,347±0,047	0,501±0,058	0,357±0,040	0,511±0,071

Tabla A.2: Error cuadrático medio para el regresor lineal en el caso del experimento tipo 1 para cada combinación de LEDs 2.

RMSE – Coord.	LEDs					
	12	13	14	23	24	34
Coord. X (m)	0,362±0,048	0,434±0,041	0,193±0,037	0,257±0,07	0,542±0,053	0,354±0,041
Coord. Y (m)	0,165±0,027	0,302±0,032	0,303±0,050	0,297±0,038	0,472±0,053	0,215±0,045

Tabla A.3: Error cuadrático medio para el regresor lineal en el caso del experimento tipo 2 para cada combinación de LEDs 1.

RMSE – Dist.(m)	LEDs			
	1	2	3	4
LED 1	0,238±0,036	–	–	–
LED 2	–	0,174±0,022	–	–
LED 3	–	–	0,204±0,028	–
LED 4	–	–	–	0,146±0,015

Tabla A.4: Error cuadrático medio para el regresor lineal en el caso del experimento tipo 2 para cada combinación de LEDs 2.

RMSE – Dist.(m)	LEDs					
	12	13	14	23	24	34
LED 1	0,217±0,033	0,182±0,037	0,230±0,029	–	–	–
LED 2	0,154±0,020	–	–	0,170±0,022	0,167±0,026	–
LED 3	–	0,147±0,020	–	0,196±0,023	–	0,198±0,025
LED 4	–	–	0,147±0,015	–	0,133±0,013	0,140±0,014

Tabla A.5: Error medio absoluto para el regresor lineal en el caso del experimento tipo 1 para cada combinación de LEDs 1.

MAE – Coord	LEDs			
	1	2	3	4
Coord. X (m)	0,380±0,040	0,443±0,047	0,383±0,043	0,446±0,046
Coord. Y (m)	0,282±0,037	0,409±0,046	0,295±0,035	0,420±0,048

Tabla A.6: Error medio absoluto para el regresor lineal en el caso del experimento tipo 1 para cada combinación de LEDs 2.

MAE - Coord.	LEDs					
	12	13	14	23	24	34
Coord. X (m)	0,282±0,040	0,375±0,040	0,131±0,023	0,157±0,032	0,439±0,047	0,278±0,036
Coord. Y (m)	0,129±0,016	0,257±0,029	0,239±0,034	0,233±0,028	0,390±0,043	0,148±0,024

Tabla A.7: Error medio absoluto para el regresor lineal en el caso del experimento tipo 2 para cada combinación de LEDs 1.

MAE – Dist.(m)	LEDs			
	1	2	3	4
LED 1	0,185±0,020	–	–	–
LED 2	–	0,141±0,015	–	–
LED 3	–	–	0,168±0,017	–
LED 4	–	–	–	0,119±0,013

Tabla A.8: Error medio absoluto para el regresor lineal en el caso del experimento tipo 2 para cada combinación de LEDs 2.

MAE – Dist.(m)	LEDs					
	12	13	14	23	24	34
LED 1	0,166±0,017	0,128±0,018	0,181±0,021	–	–	–
LED 2	0,122±0,013	–	–	0,135±0,015	0,129±0,015	–
LED 3	–	0,120±0,012	–	0,163±0,015	–	0,162±0,018
LED 4	–	–	0,120±0,013	–	0,112±0,012	0,115±0,012

Tabla A.9: Error relativo medio para el regresor lineal en el caso del experimento tipo 1 para cada combinación de LEDs 1.

MAER – Coord.	LEDs			
	1	2	3	4
Coord. X (m)	0,293±0,060	0,375±0,091	0,318±0,074	0,346±0,073
Coord. Y (m)	0,229±0,046	0,367±0,085	0,225±0,045	0,323±0,058

Tabla A.10: Error relativo medio para el regresor lineal en el caso del experimento tipo 1 para cada combinación de LEDs 2.

MAER – Coord.	LEDs					
	12	13	14	23	24	34
Coord. X (m)	0,238±0,064	0,297±0,065	0,079±0,015	0,182±0,069	0,352±0,081	0,221±0,055
Coord. Y (m)	0,102±0,024	0,194±0,034	0,188±0,037	0,196±0,049	0,324±0,071	0,106±0,022

Tabla A.11: Error relativo medio para el regresor lineal en el caso del experimento tipo 2 para cada combinación de LEDs 1.

MAER – Dist.(m)	LEDs				
	1	2	3	4	12
LED 1	0,278±0,186	–	–	–	0,220±0,127
LED 2	–	0,164±0,073	–	–	0,135±0,055
LED 3	–	–	0,199±0,060	–	–
LED 4	–	–	–	0,117±0,033	–

Tabla A.12: Error relativo medio para el regresor lineal en el caso del experimento tipo 2 para cada combinación de LEDs 2.

MAER – Dist.(m)	LEDs					
	12	13	14	23	24	34
LED 1	0,220±0,127	0,113±0,029	0,279±0,190	–	–	–
LED 2	0,135±0,055	–	–	0,156±0,069	0,144±0,062	–
LED 3	–	0,119±0,020	–	0,198±0,062	–	0,186±0,052
LED 4	–	–	0,117±0,033	–	0,105±0,027	0,110±0,029

Tabla A.13: Error medio absoluto para el regresor lineal en el caso del experimento tipo 1 para cada combinación de LEDs 3.

MAE - Coord.	LEDs				
	123	124	134	234	1234
Coord. X (m)	0,172±0,028	0,135±0,021	0,138±0,020	0,163±0,026	0,132±0,019
Coord. Y (m)	0,117±0,016	0,130±0,017	0,144±0,020	0,150±0,023	0,110±0,015

Tabla A.14: Error medio absoluto para el regresor lineal en el caso del experimento tipo 2 para cada combinación de LEDs 3.

MAE – Dist.(m)	LEDs				
	123	124	134	234	1234
LED 1	0,129±0,016	0,130±0,014	0,130±0,018	–	0,125±0,014
LED 2	0,121±0,013	0,122±0,013	–	0,125±0,014	0,122±0,014
LED 3	0,115±0,012	–	0,120±0,011	0,133±0,014	0,115±0,011
LED 4	–	0,112±0,012	0,110±0,012	0,111±0,011	0,110±0,011

Tabla A.15: Error medio absoluto para el regresor lineal en el caso del experimento tipo 2 para cada combinación de LEDs.

MAE – Dist.(m) (Coord.)	LEDs				
	123	124	134	234	1234
Coord. X (m)	0,151±0,028	0,104±0,013	0,117±0,017	0,095±0,011	0,090±0,010
Coord. Y (m)	0,126±0,015	0,125±0,015	0,142±0,022	0,128±0,015	0,115±0,015

Tabla A.16: Error relativo medio para el regresor lineal en el caso del experimento tipo 1 para cada combinación de LEDs 3.

MAER – Coord.	LEDs				
	123	124	134	234	1234
Coord. X (m)	0,183±0,065	0,089±0,019	0,087±0,017	0,170±0,062	0,107±0,034
Coord. Y (m)	0,093±0,023	0,103±0,024	0,105±0,019	0,114±0,028	0,086±0,021

Tabla A.17: Error relativo medio para el regresor lineal en el caso del experimento tipo 2 para cada combinación de LEDs 3.

MAER – Dist.(m)	LEDs				
	123	124	134	234	1234
LED 1	0,112±0,024	0,161±0,071	0,121±0,037	–	0,130±0,042
LED 2	0,132±0,052	0,131±0,052	–	0,136±0,056	0,132±0,052
LED 3	0,114±0,022	–	0,118±0,020	0,161±0,044	0,117±0,023
LED 4	–	0,105±0,027	0,105±0,026	0,103±0,026	0,103±0,025

Tabla A.18: Error relativo medio para el regresor lineal en el caso del experimento tipo 2 para cada combinación de LEDs.

MAER – Dist.(m) (Coord.)	LEDs				
	123	124	134	234	1234
Coord. X (m)	0,164±0,063	0,080±0,017	0,083±0,018	0,075±0,018	0,071±0,017
Coord. Y (m)	0,102±0,024	0,095±0,020	0,106±0,021	0,107±0,024	0,093±0,022

## A.2.2. Gráficas del Regresor lineal

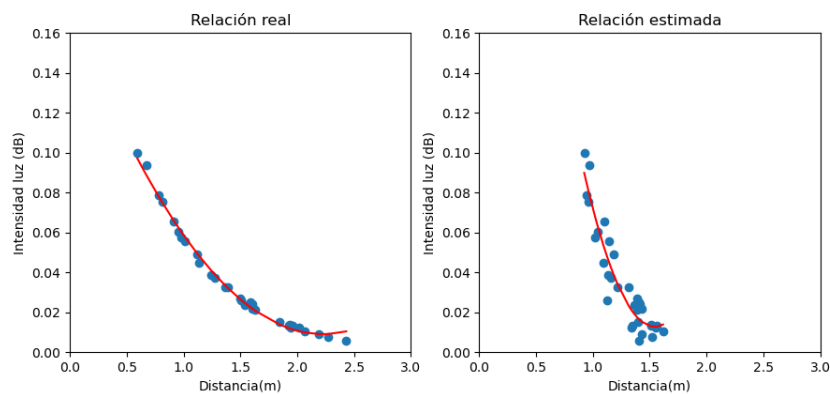


Figura A.2: Relación Intensidad-Distancia de los datos reales y los estimados para el LED 2.



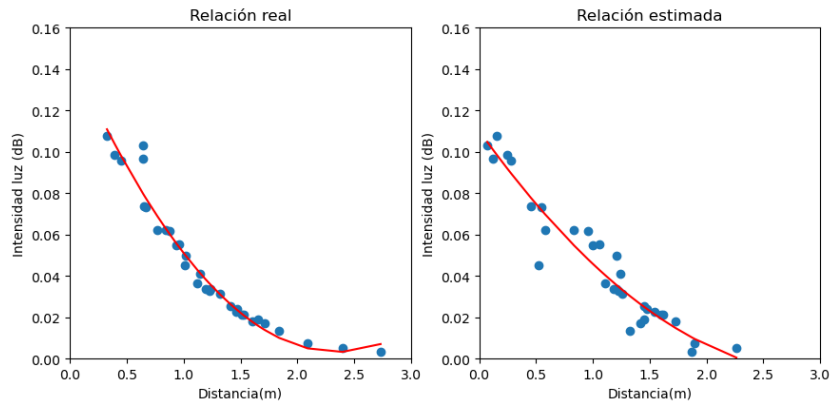


Figura A.3: Relación Intensidad-Distancia de los datos reales y los estimados para el LED 3.

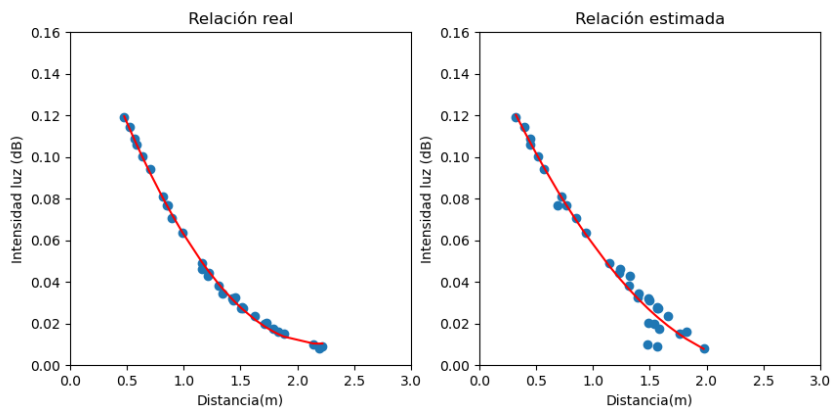


Figura A.4: Relación Intensidad-Distancia de los datos reales y los estimados para el LED 4.

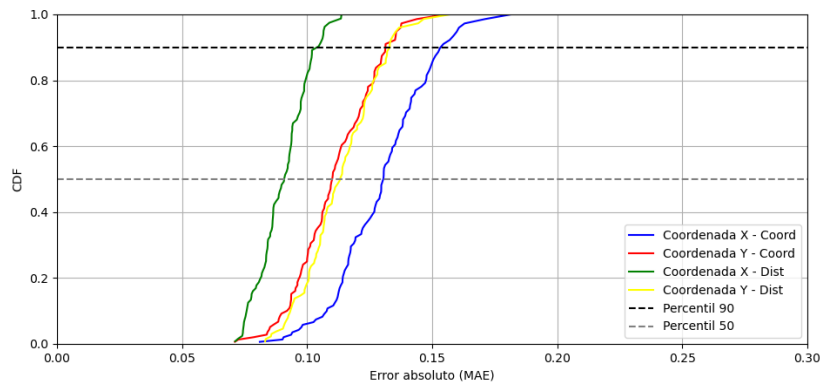


Figura A.5: CDF del MAE para el experimento tipo 1.

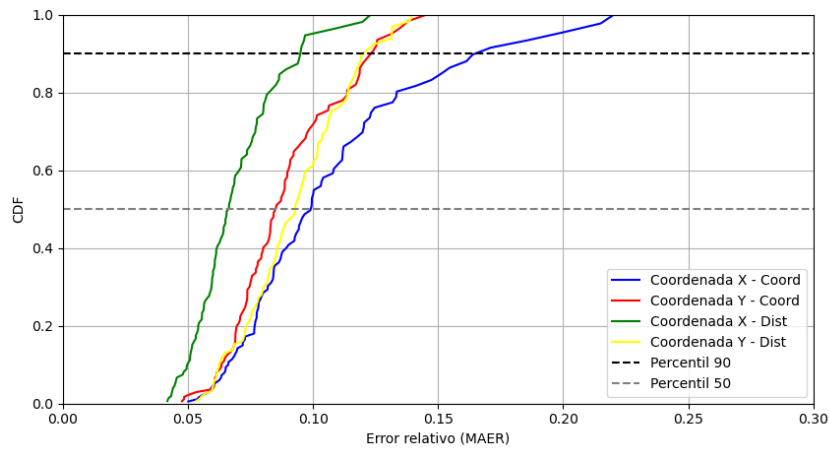


Figura A.6: CDF del MAER para el experimento tipo 1.

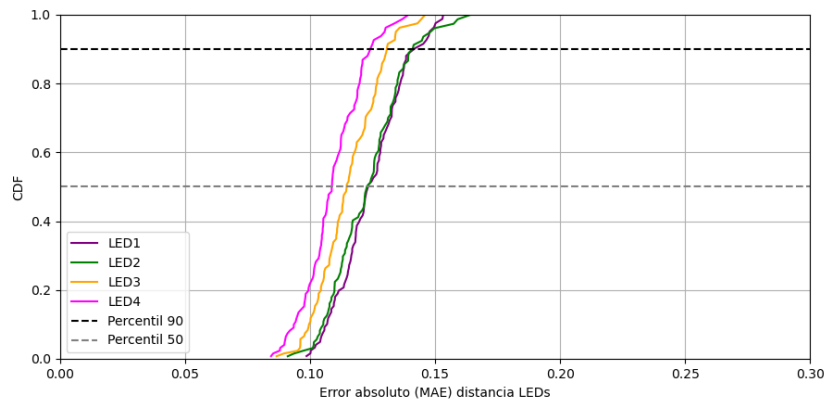


Figura A.7: CDF del MAE para el experimento tipo 2.

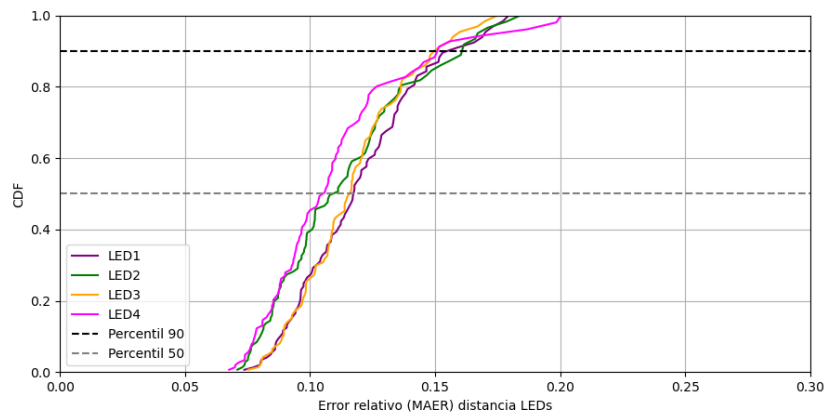


Figura A.8: CDF del MAER para el experimento tipo 2.

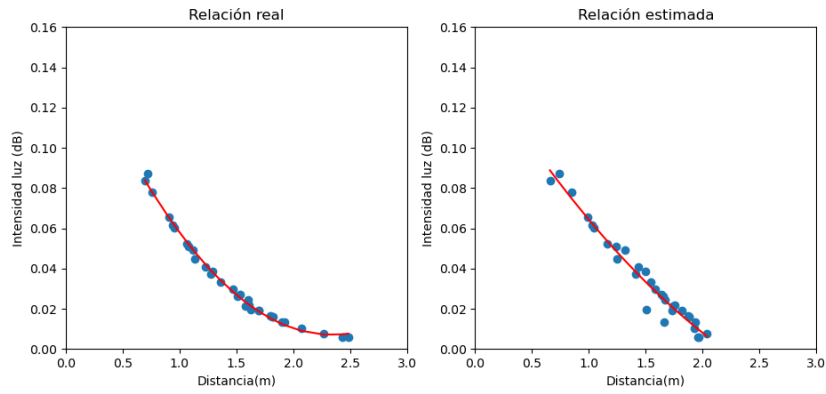


Figura A.9: Relación Intensidad-Distancia de los datos reales y los estimados para el LED 2 para el experimento tipo 2.

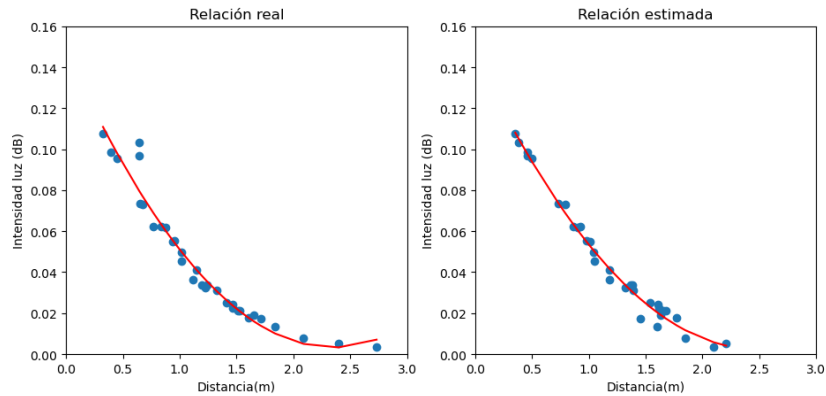


Figura A.10: Relación Intensidad-Distancia de los datos reales y los estimados para el LED 3 para el experimento tipo 2.

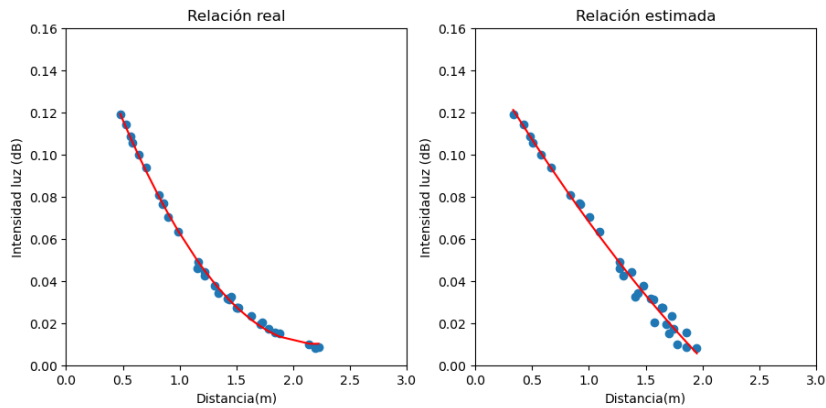
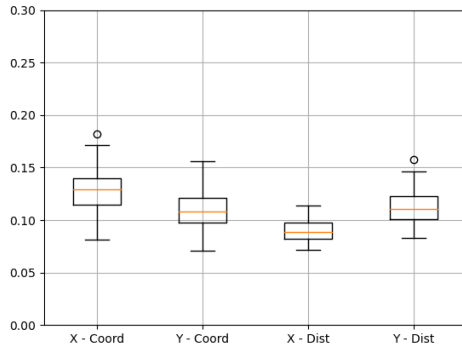
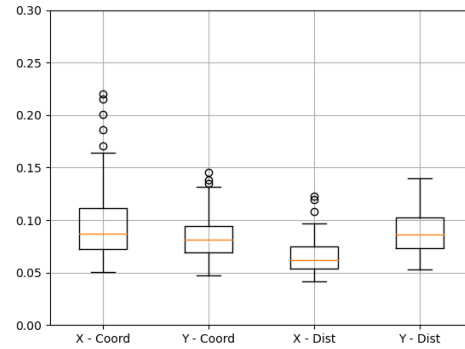


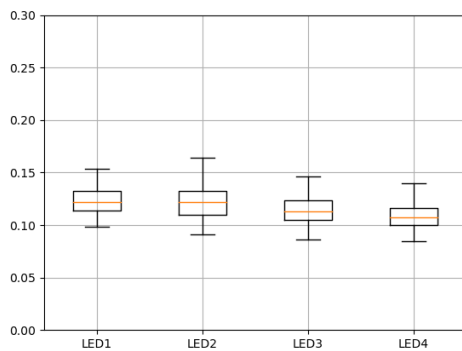
Figura A.11: Relación Intensidad-Distancia de los datos reales y los estimados para el LED 4 para el experimento tipo 2.



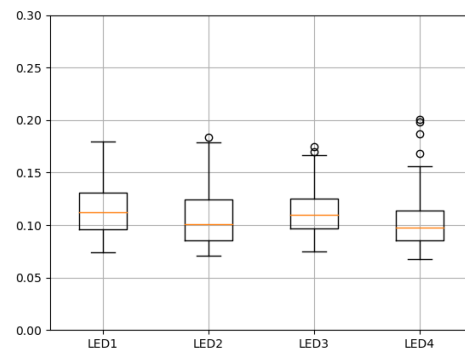
(a) Box-plot del MAE para el experimento tipo 1.



(b) Box-plot del MAER para el experimento tipo 1.



(c) Box-plot del MAE para el experimento tipo 2.



(d) Box-plot del MAER para el experimento tipo 2.

Figura A.12: Box-plot de los MAE y MAER para el regresor lineal.

### A.2.3. Tablas del Regresor SVR lineal

Tabla A.19: Error cuadrático medio para el regresor SVR lineal en el caso del experimento tipo 1 para cada combinación de LEDs 1.

RMSE – Coord.	LEDs				
	1	2	3	4	12
Coord. X (m)	0,513±0,047	0,560±0,050	0,529±0,049	0,556±0,050	0,493±0,047
Coord. Y (m)	0,517±0,059	0,609±0,061	0,568±0,058	0,602±0,063	0,470±0,059

Tabla A.20: Error cuadrático medio para el regresor SVR lineal en el caso del experimento tipo 1 para cada combinación de LEDs 2.

RMSE – Coord.	LEDs					
	12	13	14	23	24	34
Coord. X (m)	0,493±0,047	0,488±0,047	0,479±0,046	0,509±0,049	0,548±0,050	0,504±0,049
Coord. Y (m)	0,470±0,059	0,449±0,054	0,466±0,062	0,523±0,058	0,570±0,062	0,511±0,060

Tabla A.21: Error cuadrático medio para el regresor SVR lineal en el caso del experimento tipo 2 para cada combinación de LEDs 1.

RMSE – Dist.(m)	LEDs			
	1	2	3	4
LED 1	0,493±0,062	–	–	–
LED 2	–	0,431±0,042	–	–
LED 3	–	–	0,486±0,051	–
LED 4	–	–	–	0,390±0,037

Tabla A.22: Error cuadrático medio para el regresor SVR lineal en el caso del experimento tipo 2 para cada combinación de LEDs 2.

RMSE – Dist.(m)	LEDs					
	12	13	14	23	24	34
LED 1	0,485±0,060	0,405±0,059	0,488±0,065	–	–	–
LED 2	0,418±0,041	–	–	0,415±0,040	0,390±0,043	–
LED 3	–	0,355±0,041	–	0,482±0,053	–	0,478±0,049
LED 4	–	–	0,371±0,036	–	0,357±0,036	0,383±0,037

Tabla A.23: Error medio absoluto para el regresor SVR lineal en el caso del experimento tipo 1 para cada combinación de LEDs 1.

MAE – Coord.	LEDs			
	1	2	3	4
Coord. X (m)	0,418±0,043	0,458±0,046	0,429±0,045	0,456±0,047
Coord. Y (m)	0,433±0,052	0,518±0,056	0,482±0,053	0,507±0,056

Tabla A.24: Error medio absoluto para el regresor SVR lineal en el caso del experimento tipo 1 para cada combinación de LEDs 2.

MAE – Coord.	LEDs					
	12	13	14	23	24	34
Coord. X (m)	0,396±0,042	0,394±0,041	0,387±0,042	0,409±0,044	0,447±0,047	0,405±0,045
Coord. Y (m)	0,389±0,050	0,371±0,046	0,375±0,050	0,438±0,052	0,475±0,055	0,426±0,051

Tabla A.25: Error medio absoluto para el regresor SVR lineal en el caso del experimento tipo 2 para cada combinación de LEDs 1.

MAE – Dist.(m)	LEDs			
	1	2	3	4
LED 1	0,407±0,048	–	–	–
LED 2	–	0,358±0,094	–	–
LED 3	–	–	0,401±0,044	–
LED 4	–	–	–	0,326±0,036

Tabla A.26: Error medio absoluto para el regresor SVR lineal en el caso del experimento tipo 2 para cada combinación de LEDs 2.

MAE – Dist.(m)	LEDs					
	12	13	14	23	24	34
LED 1	0,400±0,047	0,329±0,043	0,393±0,048	–	–	–
LED 2	0,347±0,039	–	–	0,345±0,038	0,319±0,039	–
LED 3	–	0,288±0,035	–	0,393±0,045	–	0,394±0,043
LED 4	–	–	0,306±0,034	–	0,296±0,035	0,317±0,035

Tabla A.27: Error relativo medio para el regresor SVR lineal en el caso del experimento tipo 1 para cada combinación de LEDs 1.

MAER – Coord.	LEDs			
	1	2	3	4
Coord. X (m)	0,325±0,069	0,361±0,077	0,340±0,074	0,350±0,072
Coord. Y (m)	0,349±0,071	0,438±0,088	0,412±0,085	0,425±0,085

Tabla A.28: Error relativo medio para el regresor SVR lineal en el caso del experimento tipo 2 para cada combinación de LEDs 2.

MAER – Coord.	LEDs					
	12	13	14	23	24	34
Coord. X (m)	0,316±0,070	0,308±0,068	0,300±0,065	0,328±0,073	0,347±0,072	0,316±0,069
Coord. Y (m)	0,314±0,067	0,306±0,066	0,299±0,064	0,379±0,081	0,402±0,083	0,366±0,077

Tabla A.29: Error relativo medio para el regresor SVR lineal en el caso del experimento tipo 1 para cada combinación de LEDs 1.

MAER – Dist.(m)	LEDs			
	1	2	3	4
LED 1	0,468±0,108	–	–	–
LED 2	–	0,371±0,094	–	–
LED 3	–	–	0,497±0,125	–
LED 4	–	–	–	0,364±0,091

Tabla A.30: Error relativo medio para el regresor SVR lineal en el caso del experimento tipo 2 para cada combinación de LEDs 2.

MAER – Dist.(m)	LEDs					
	12	13	14	23	24	34
LED 1	0,469±0,110	0,390±0,094	0,438±0,101	–	–	–
LED 2	0,365±0,095	–	–	0,354±0,090	0,331±0,087	–
LED 3	–	0,390±0,109	–	0,483±0,122	–	0,503±0,128
LED 4	–	–	0,336±0,084	–	0,330±0,083	0,359±0,091

Tabla A.31: Error medio absoluto para el regresor SVR lineal en el caso del experimento tipo 1 para cada combinación de LEDs 3.

MAE – Coord.	LEDs				
	123	124	134	234	1234
Coord. X (m)	0,366±0,040	0,369±0,042	0,355±0,041	0,393±0,044	0,332±0,040
Coord. Y (m)	0,323±0,045	0,338±0,049	0,309±0,045	0,389±0,050	0,269±0,044

Tabla A.32: Error medio absoluto para el regresor SVR lineal en el caso del experimento tipo 2 para cada combinación de LEDs 3.

MAE – Dist.(m)	LEDs				
	123	124	134	234	1234
LED 1	0,330±0,042	0,392±0,047	0,321±0,043	–	0,323±0,043
LED 2	0,340±0,037	0,312±0,038	–	0,306±0,037	0,304±0,037
LED 3	0,283±0,036	–	0,287±0,036	0,390±0,044	0,283±0,036
LED 4	–	0,278±0,033	0,305±0,034	0,292±0,034	0,279±0,033

Tabla A.33: Error medio absoluto para el regresor SVR lineal en el caso del experimento tipo 2 para cada combinación de LEDs.

MAE – Dist.(m) (Coord.)	LEDs				
	123	124	134	234	1234
Coord. X (m)	0,378±0,047	0,309±0,045	0,360±0,042	0,361±0,054	0,278±0,040
Coord. Y (m)	0,348±0,043	0,319±0,046	0,318±0,051	0,353±0,053	0,255±0,045

Tabla A.34: Error relativo medio para el regresor SVR lineal en el caso del experimento tipo 1 para cada combinación de LEDs 3.

MAER – Coord.	LEDs				
	123	124	134	234	1234
Coord. X (m)	0,293±0,068	0,292±0,065	0,278±0,063	0,308±0,069	0,265±0,063
Coord. Y (m)	0,269±0,061	0,270±0,060	0,254±0,058	0,338±0,074	0,222±0,055

Tabla A.35: Error relativo medio para el regresor SVR lineal en el caso del experimento tipo 1 para cada combinación de LEDs 3.

MAER – Dist.(m)	LEDs				
	123	124	134	234	1234
LED 1	0,403±0,100	0,451±0,104	0,372±0,089	–	0,384±0,096
LED 2	0,353±0,092	0,327±0,088	–	0,315±0,083	0,315±0,085
LED 3	0,378±0,108	–	0,392±0,112	0,494±0,126	0,382±0,108
LED 4	–	0,308±0,080	0,339±0,086	0,330±0,085	0,311±0,082

Tabla A.36: Error relativo medio para el regresor SVR lineal en el caso del experimento tipo 2 para cada combinación de LEDs.

MAER – Dist.(m) (Coord.)	LEDs				
	123	124	134	234	1234
Coord. X (m)	$0,327 \pm 0,078$	$0,245 \pm 0,058$	$0,248 \pm 0,049$	$0,302 \pm 0,084$	$0,228 \pm 0,062$
Coord. Y (m)	$0,316 \pm 0,071$	$0,264 \pm 0,063$	$0,235 \pm 0,051$	$0,290 \pm 0,066$	$0,217 \pm 0,058$

#### A.2.4. Gráficas del Regresor SVR lineal

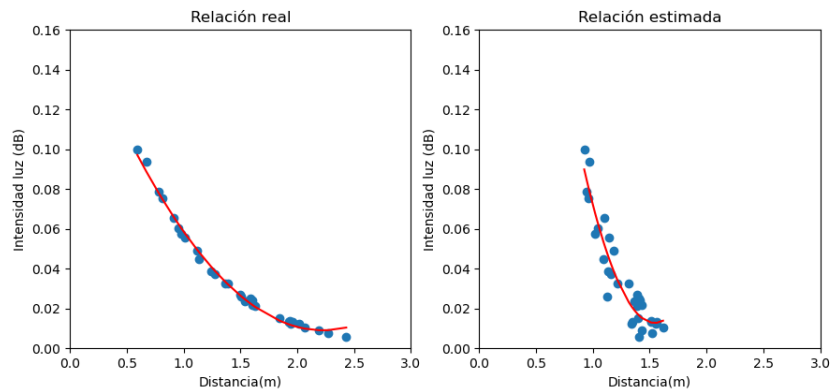


Figura A.13: Relación Intensidad-Distancia de los datos reales y los estimados para el LED 2.

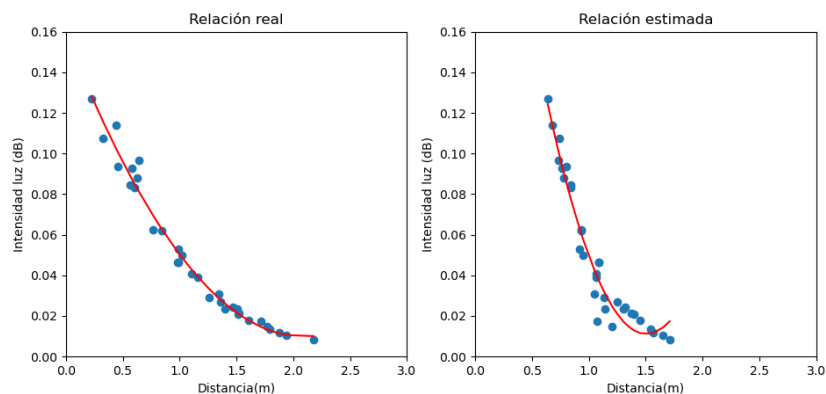


Figura A.14: Relación Intensidad-Distancia de los datos reales y los estimados para el LED 3.



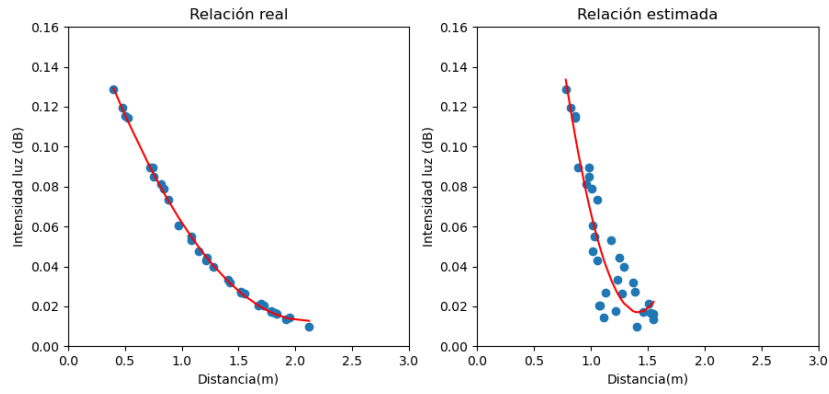


Figura A.15: Relación Intensidad-Distancia de los datos reales y los estimados para el LED 4.

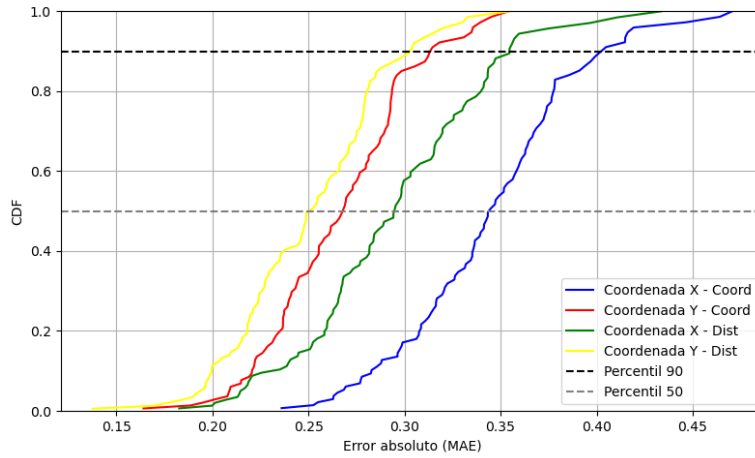


Figura A.16: CDF del MAE para el experimento tipo 1.

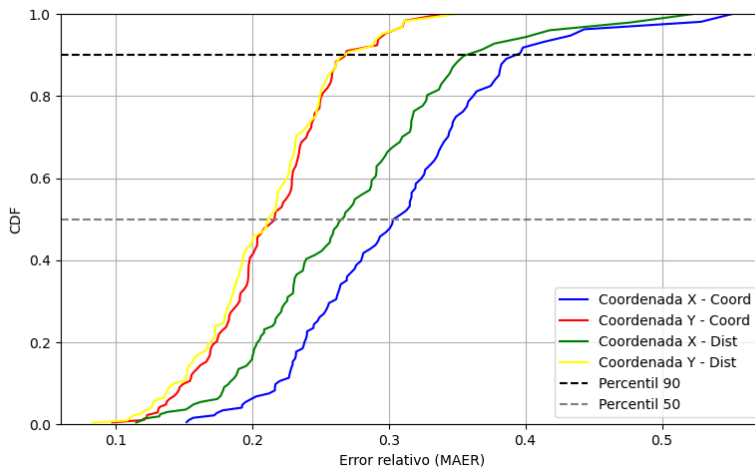


Figura A.17: CDF del MAER para el experimento tipo 1.

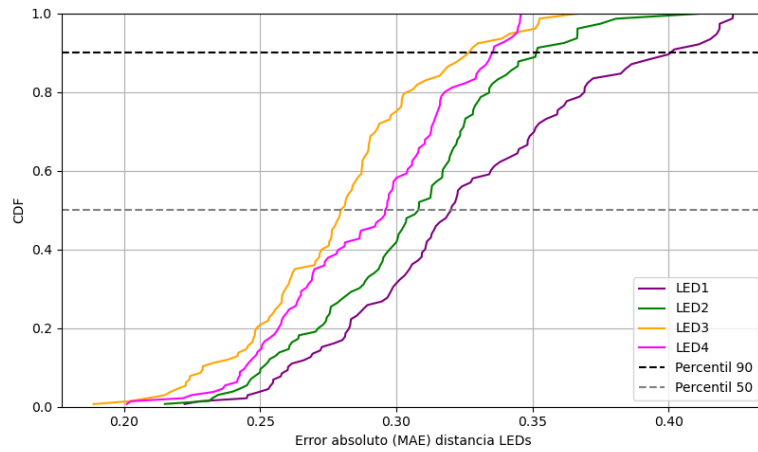


Figura A.18: CDF del MAE para el experimento tipo 2.

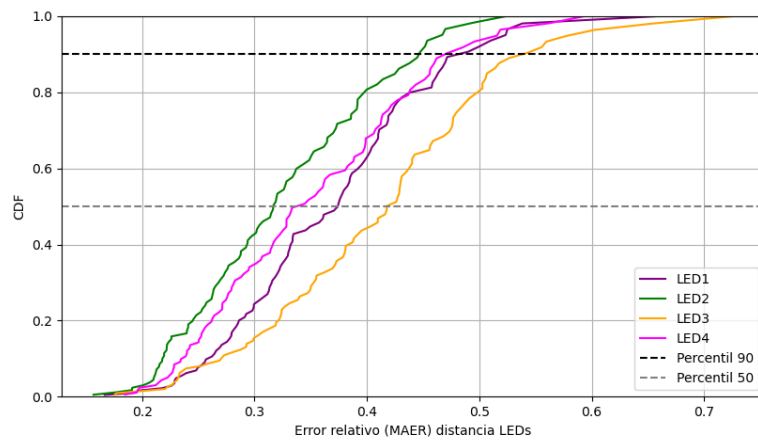


Figura A.19: CDF del MAER para el experimento tipo 2.

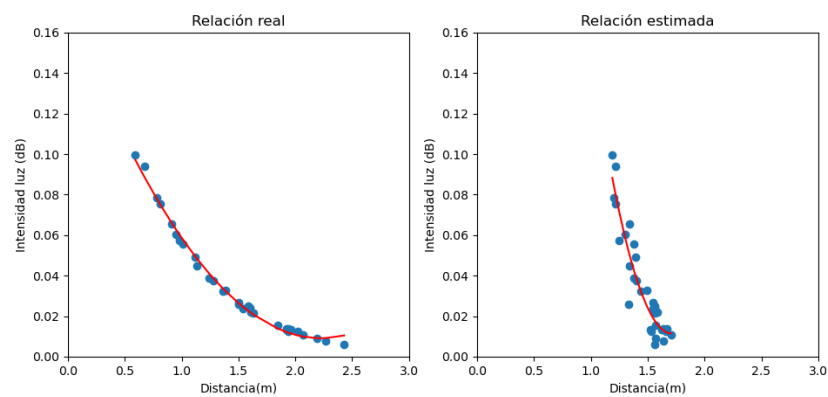


Figura A.20: Relación Intensidad-Distancia de los datos reales y los estimados para el LED 2 para el experimento tipo 2.

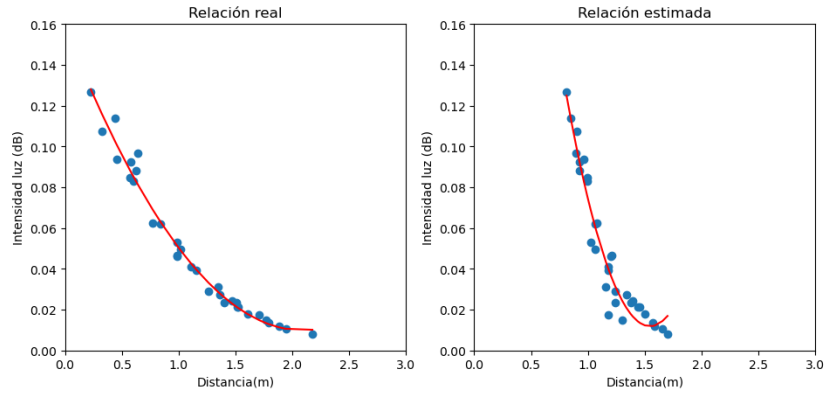


Figura A.21: Relación Intensidad-Distancia de los datos reales y los estimados para el LED 3 para el experimento tipo 2.

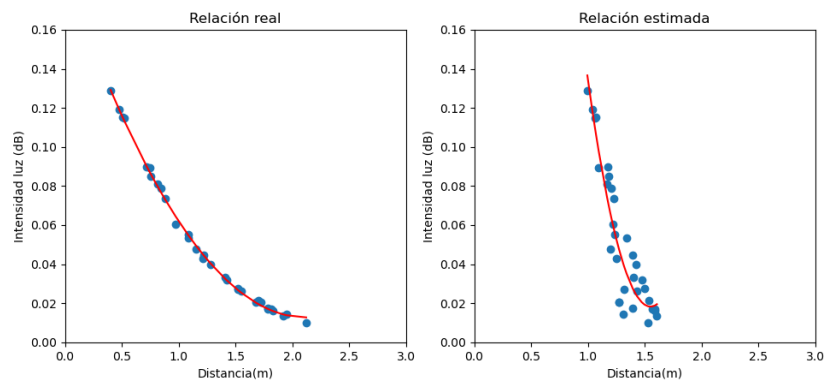
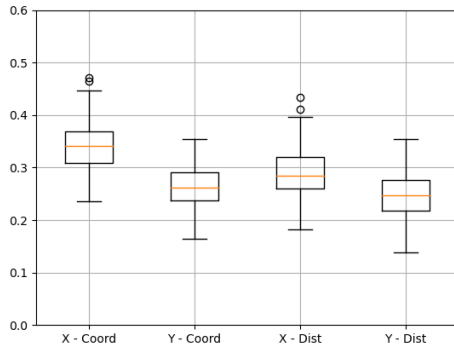
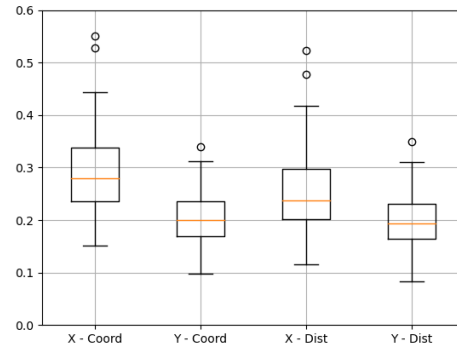


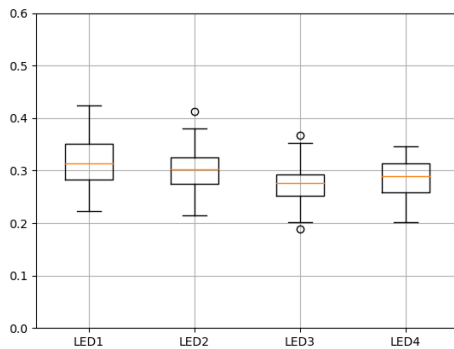
Figura A.22: Relación Intensidad-Distancia de los datos reales y los estimados para el LED 4 para el experimento tipo 2.



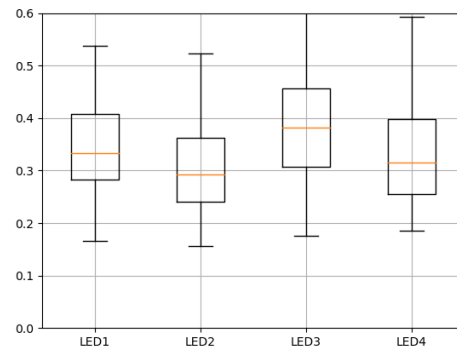
(a) Box-plot del MAE para el experimento tipo 1.



(b) Box-plot del MAER para el experimento tipo 1.



(c) Box-plot del MAE para el experimento tipo 2.



(d) Box-plot del MAER para el experimento tipo 2.

Figura A.23: Box-plot de los MAE y MAER para el regresor SVR lineal.

### A.2.5. Tablas del Regresor SVR polinomial

Tabla A.37: Error cuadrático medio para el regresor SVR polinomial en el caso del experimento tipo 1 para cada combinación de LEDs 1.

RMSE – Coord.	LEDs				
	1	2	3	4	12
Coord. X (m)	0,504±0,050	0,557±0,065	0,515±0,056	0,562±0,055	0,481±0,057
Coord. Y (m)	0,475±0,055	0,574±0,054	0,489±0,051	0,561±0,073	0,324±0,048

Tabla A.38: Error cuadrático medio para el regresor SVR polinomial en el caso del experimento tipo 1 para cada combinación de LEDs 2.

RMSE – Coord.	LEDs					
	12	13	14	23	24	34
Coord. X (m)	0,481±0,057	0,491±0,056	0,260±0,051	0,333±0,085	0,566±0,061	0,485±0,053
Coord. Y (m)	0,324±0,048	0,376±0,057	0,438±0,061	0,464±0,053	0,520±0,063	0,354±0,056

Tabla A.39: Error cuadrático medio para el regresor SVR polinomial en el caso del experimento tipo 2 para cada combinación de LEDs 1.

RMSE – Dist.(m)	LEDs				
	1	2	3	4	12
LED 1	0,445±0,059	–	–	–	0,391±0,057
LED 2	–	0,370±0,046	–	–	0,303±0,035
LED 3	–	–	0,408±0,044	–	–
LED 4	–	–	–	0,338±0,048	–

Tabla A.40: Error cuadrático medio para el regresor SVR polinomial en el caso del experimento tipo 2 para cada combinación de LEDs 2.

RMSE – Dist.(m)	LEDs					
	12	13	14	23	24	34
LED 1	0,391±0,057	0,260±0,057	0,349±0,049	–	–	–
LED 2	0,303±0,035	–	–	0,272±0,032	0,199±0,037	–
LED 3	–	0,200±0,034	–	0,326±0,040	–	0,366±0,041
LED 4	–	–	0,245±0,035	–	0,189±0,027	0,285±0,031

Tabla A.41: Error medio absoluto para el regresor SVR polinomial en el caso del experimento tipo 1 para cada combinación de LEDs 1.

MAE – Coord.	LEDs				
	1	2	3	4	12
Coord. X (m)	0,406±0,045	0,431±0,055	0,399±0,047	0,453±0,051	0,371±0,047
Coord. Y (m)	0,381±0,047	0,464±0,050	0,397±0,044	0,430±0,057	0,251±0,031

Tabla A.42: Error medio absoluto para el regresor SVR polinomial en el caso del experimento tipo 1 para cada combinación de LEDs 2.

MAE – Coord.	LEDs					
	12	13	14	23	24	34
Coord. X (m)	0,371±0,047	0,384±0,042	0,188±0,027	0,210±0,043	0,437±0,053	0,373±0,045
Coord. Y (m)	0,251±0,031	0,277±0,040	0,313±0,048	0,359±0,045	0,387±0,052	0,261±0,038

Tabla A.43: Error medio absoluto para el regresor SVR polinomial en el caso del experimento tipo 2 para cada combinación de LEDs 1.

MAE – Dist.(m)	LEDs			
	1	2	3	4
LED 1	0,348±0,041	–	–	–
LED 2	–	0,298±0,034	–	–
LED 3	–	–	0,353±0,035	–
LED 4	–	–	–	0,298±0,033

Tabla A.44: Error medio absoluto para el regresor SVR polinomial en el caso del experimento tipo 2 para cada combinación de LEDs 2.

MAE – Dist.(m)	LEDs					
	12	13	14	23	24	34
LED 1	0,298±0,038	0,173±0,030	0,239±0,036	–	–	–
LED 2	0,220±0,030	–	–	0,183±0,027	0,120±0,023	–
LED 3	–	0,156±0,021	–	0,237±0,030	–	0,305±0,034
LED 4	–	–	0,179±0,029	–	0,135±0,021	0,242±0,027

Tabla A.45: Error relativo medio para el regresor SVR polinomial en el caso del experimento tipo 1 para cada combinación de LEDs 1.

MAER – Coord.	LEDs				
	1	2	3	4	12
Coord. X (m)	0,314±0,069	0,370±0,095	0,338±0,083	0,340±0,074	0,314±0,082
Coord. Y (m)	0,307±0,067	0,410±0,081	0,328±0,060	0,340±0,062	0,179±0,031

Tabla A.46: Error relativo medio para el regresor SVR polinomial en el caso del experimento tipo 1 para cada combinación de LEDs 2.

MAER – Coord.	LEDs					
	12	13	14	23	24	34
Coord. X (m)	0,314±0,082	0,318±0,083	0,151±0,044	0,209±0,089	0,341±0,093	0,297±0,073
Coord. Y (m)	0,179±0,031	0,241±0,064	0,263±0,070	0,313±0,069	0,349±0,085	0,230±0,053

Tabla A.47: Error relativo medio para el regresor SVR polinomial en el caso del experimento tipo 2 para cada combinación de LEDs 1.

MAER – Dist.(m)	LEDs			
	1	2	3	4
LED 1	0,355±0,074	–	–	–
LED 2	–	0,304±0,096	–	–
LED 3	–	–	0,328±0,099	–
LED 4	–	–	–	0,265±0,120

Tabla A.48: Error relativo medio para el regresor SVR polinomial en el caso del experimento tipo 2 para cada combinación de LEDs 2.

MAER – Dist.(m)	LEDs					
	12	13	14	23	24	34
LED 1	0,305±0,059	0,172±0,033	0,248±0,045	–	–	–
LED 2	0,238±0,054	–	–	0,209±0,035	0,142±0,022	–
LED 3	–	0,146±0,034	–	0,245±0,051	–	0,285±0,070
LED 4	–	–	0,177±0,053	–	0,140±0,030	0,227±0,070

Tabla A.49: Error medio absoluto para el regresor SVR polinomial en el caso del experimento tipo 1 para cada combinación de LEDs 3.

MAE – Coord.	LEDs				
	123	124	134	234	1234
Coord. X (m)	0,181±0,041	0,173±0,028	0,161±0,027	0,197±0,041	0,011±0,042
Coord. Y (m)	0,149±0,028	0,158±0,029	0,168±0,035	0,184±0,042	0,004±0,035

Tabla A.50: Error medio absoluto para el regresor SVR polinomial en el caso del experimento tipo 2 para cada combinación de LEDs 3.

MAE – Dist.(m)	LEDs				
	123	124	134	234	1234
LED 1	0,147±0,030	0,158±0,021	0,152±0,026	–	0,114±0,019
LED 2	0,081±0,014	0,098±0,020	–	0,103±0,020	0,078±0,012
LED 3	0,118±0,021	–	0,130±0,019	0,161±0,026	0,092±0,015
LED 4	–	0,110±0,018	0,088±0,014	0,101±0,019	0,068±0,012

Tabla A.51: Error medio absoluto para el regresor SVR polinomial en el caso del experimento tipo 2 para cada combinación de LEDs.

MAE – Dist.(m) (Coord.)	LEDs				
	123	124	134	234	1234
Coord. X (m)	0,161±0,043	0,122±0,028	0,139±0,026	0,120±0,027	0,070±0,013
Coord. Y (m)	0,141±0,029	0,131±0,027	0,144±0,038	0,135±0,029	0,088±0,017

Tabla A.52: Error relativo medio para el regresor SVR polinomial en el caso del experimento tipo 1 para cada combinación de LEDs 3.

MAER – Coord.	LEDs				
	123	124	134	234	1234
Coord. X (m)	0,188±0,086	0,138±0,044	0,122±0,039	0,205±0,090	0,107±0,043
Coord. Y (m)	0,140±0,047	0,148±0,048	0,137±0,040	0,144±0,037	0,115±0,038

Tabla A.53: Error relativo medio para el regresor SVR polinomial en el caso del experimento tipo 2 para cada combinación de LEDs 3.

MAER – Dist.(m)	LEDs				
	123	124	134	234	1234
LED 1	0,143±0,041	0,132±0,042	0,155±0,029	–	0,114±0,043
LED 2	0,091±0,021	0,113±0,024	–	0,121±0,019	0,078±0,016
LED 3	0,120±0,026	–	0,122±0,033	0,134±0,051	0,092±0,022
LED 4	–	0,118±0,023	0,085±0,029	0,109±0,029	0,068±0,020

Tabla A.54: Error relativo medio para el regresor SVR polinomial en el caso del experimento tipo 2 para cada combinación de LEDs.

MAER – Dist.(m) (Coord.)	LEDs				
	123	124	134	234	1234
Coord. X (m)	$0,185 \pm 0,095$	$0,103 \pm 0,036$	$0,099 \pm 0,028$	$0,106 \pm 0,064$	$0,055 \pm 0,020$
Coord. Y (m)	$0,138 \pm 0,051$	$0,115 \pm 0,046$	$0,098 \pm 0,028$	$0,108 \pm 0,029$	$0,078 \pm 0,028$

### A.2.6. Gráficas del Regresor SVR polinomial

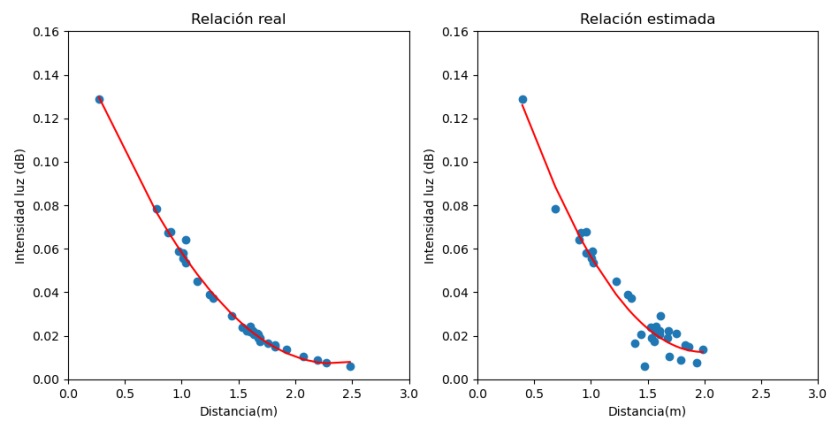


Figura A.24: Relación Intensidad-Distancia de los datos reales y los estimados para el LED 2.

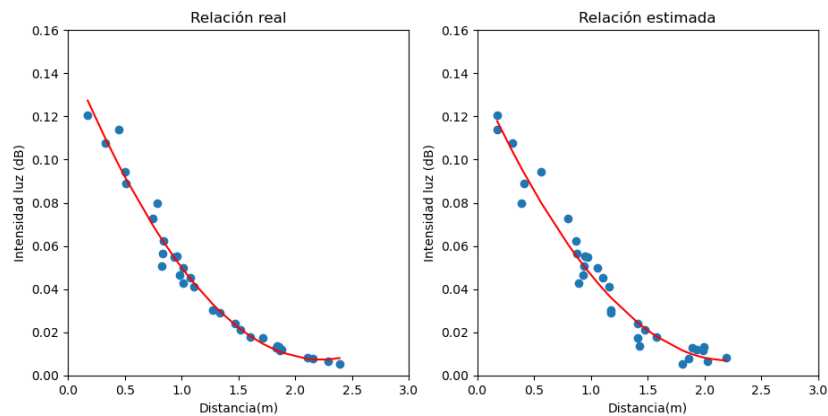


Figura A.25: Relación Intensidad-Distancia de los datos reales y los estimados para el LED 3.



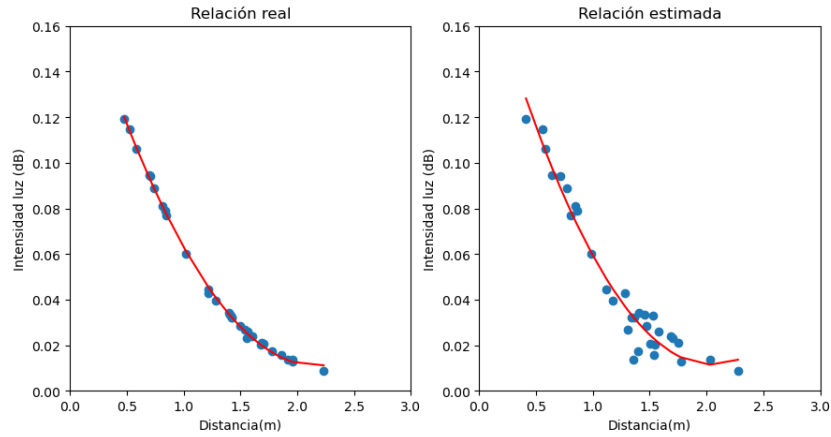


Figura A.26: Relación Intensidad-Distancia de los datos reales y los estimados para el LED 4.

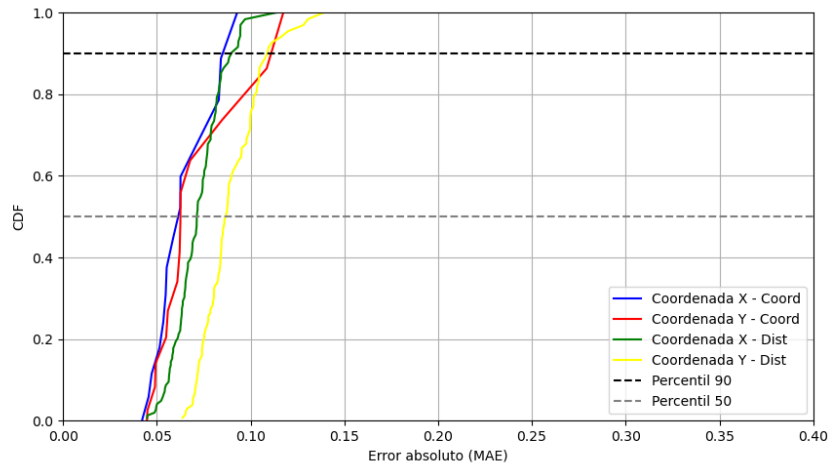


Figura A.27: CDF del MAE para el experimento tipo 2.

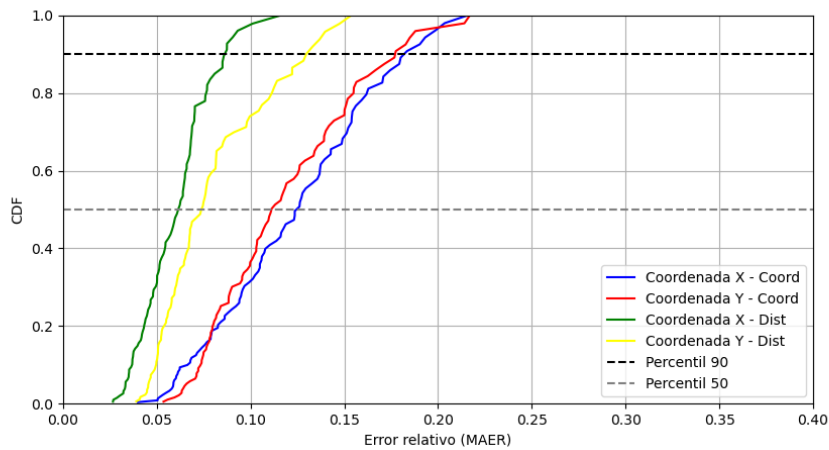


Figura A.28: CDF del MAER para el experimento tipo 1.

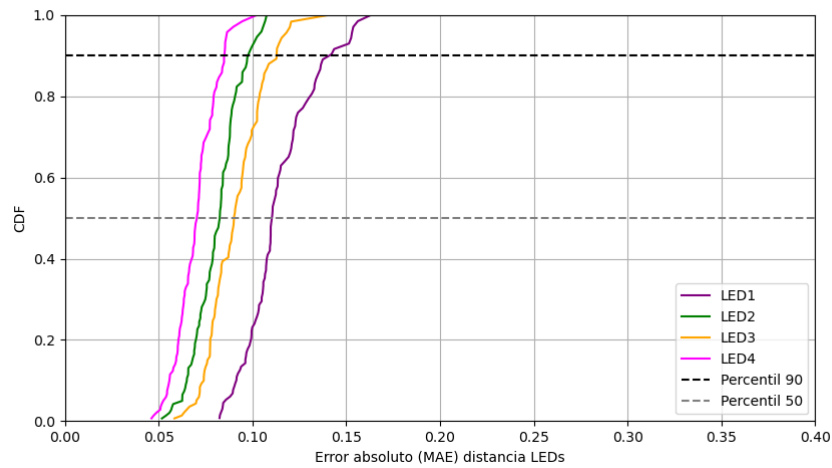


Figura A.29: CDF del MAE para el experimento tipo 2.

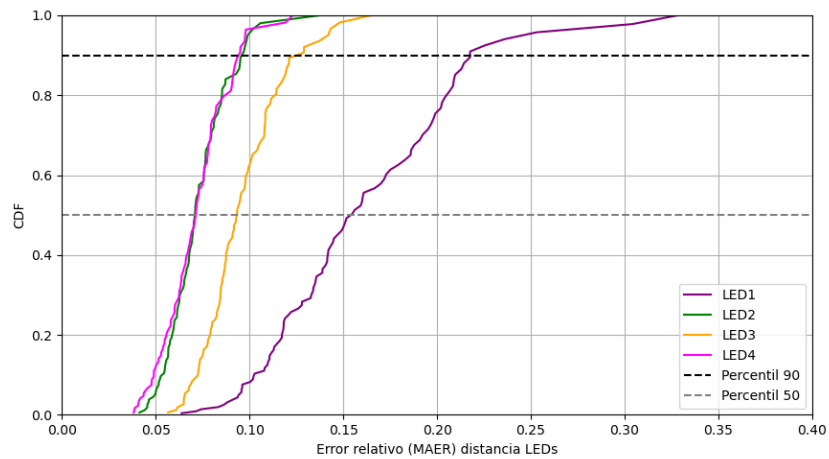


Figura A.30: CDF del MAER para el experimento tipo 2.

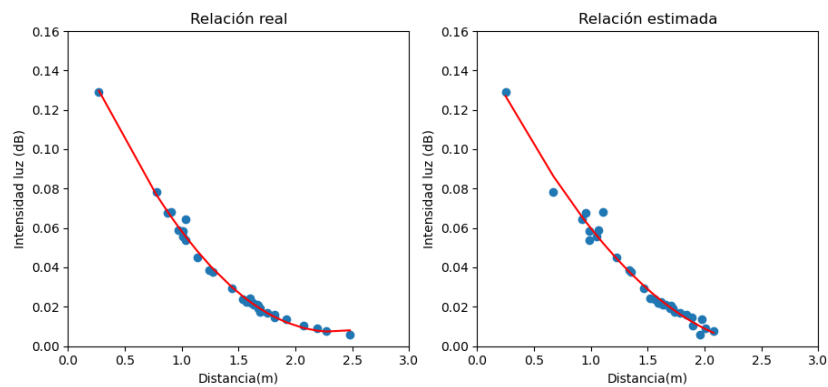


Figura A.31: Relación Intensidad-Distancia de los datos reales y los estimados para el LED 2 para el experimento tipo 2.

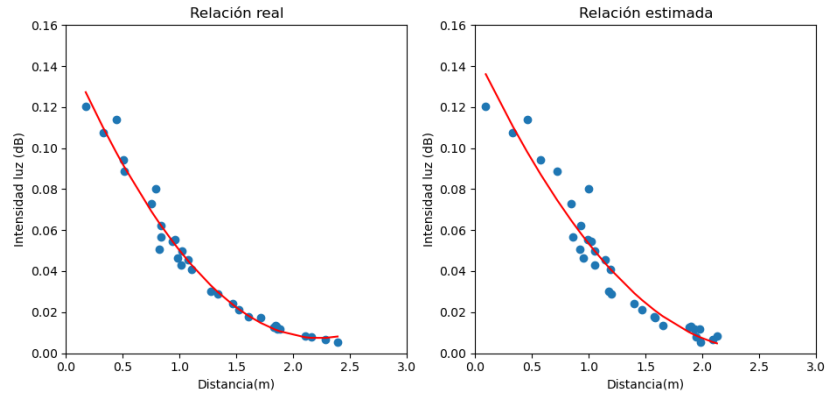


Figura A.32: Relación Intensidad-Distancia de los datos reales y los estimados para el LED 3 para el experimento tipo 2.

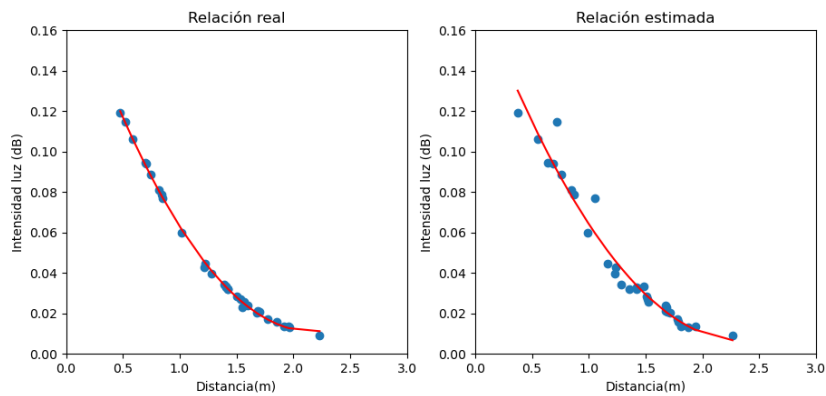
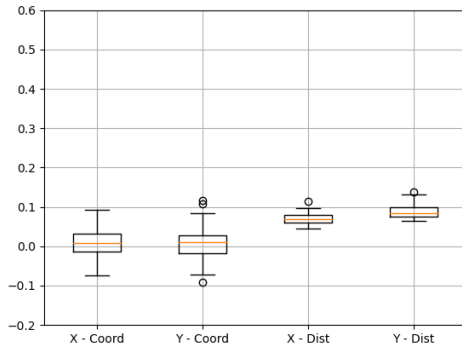
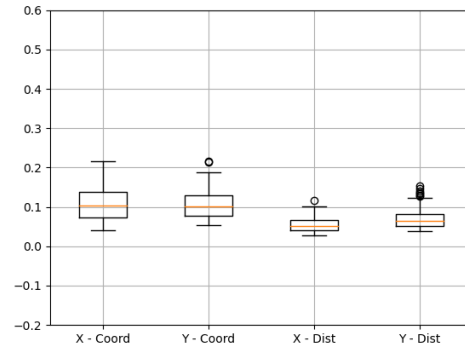


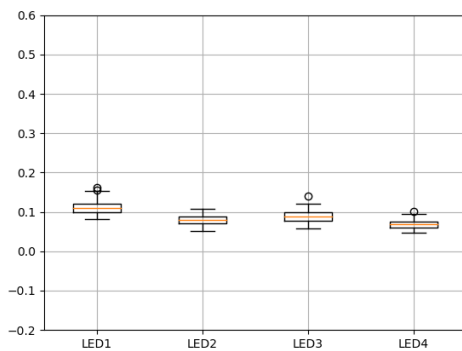
Figura A.33: Relación Intensidad-Distancia de los datos reales y los estimados para el LED 4 para el experimento tipo 2.



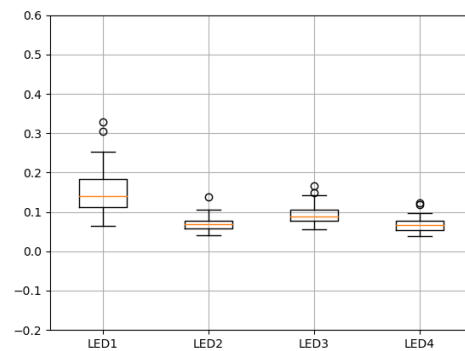
(a) Box-plot del MAE para el experimento tipo 1.



(b) Box-plot del MAER para el experimento tipo 1.



(c) Box-plot del MAE para el experimento tipo 2.



(d) Box-plot del MAER para el experimento tipo 2.

Figura A.34: Box-plot de los MAE y MAER para el regresor SVR polinomial.

### A.2.7. Tablas del Regresor SVR rbf

Tabla A.55: Error cuadrático medio para el regresor SVR rbf en el caso del experimento tipo 1 para cada combinación de LEDs 1.

RMSE – Coord.	LEDs			
	1	2	3	4
Coord. X (m)	$0,429 \pm 0,045$	$0,543 \pm 0,065$	$0,447 \pm 0,064$	$0,535 \pm 0,057$
Coord. Y (m)	$0,314 \pm 0,044$	$0,46 \pm 0,082$	$0,337 \pm 0,053$	$0,534 \pm 0,103$

Tabla A.56: Error cuadrático medio para el regresor SVR rbf en el caso del experimento tipo 1 para cada combinación de LEDs 2.

RMSE – Coord.	LEDs					
	12	13	14	23	24	34
Coord. X (m)	$0,164 \pm 0,041$	$0,480 \pm 0,075$	$0,166 \pm 0,034$	$0,235 \pm 0,078$	$0,572 \pm 0,086$	$0,144 \pm 0,034$
Coord. Y (m)	$0,113 \pm 0,026$	$0,331 \pm 0,048$	$0,154 \pm 0,043$	$0,137 \pm 0,029$	$0,497 \pm 0,079$	$0,178 \pm 0,061$

Tabla A.57: Error cuadrático medio para el regresor SVR rbf en el caso del experimento tipo 2 para cada combinación de LEDs 1.

RMSE – Dist.(m)	LEDs			
	1	2	3	4
LED 1	0,098±0,026	–	–	–
LED 2	–	0,068±0,01	–	–
LED 3	–	–	0,081±0,016	–
LED 4	–	–	–	0,059±0,008

Tabla A.58: Error cuadrático medio para el regresor SVR rbf en el caso del experimento tipo 2 para cada combinación de LEDs 2.

RMSE – Dist.(m)	LEDs					
	12	13	14	23	24	34
LED 1	0,101±0,027	0,106±0,031	0,117±0,031	–	–	–
LED 2	0,081±0,013	–	–	0,084±0,015	0,096±0,026	–
LED 3	–	0,089±0,021	–	0,088±0,02	–	0,092±0,019
LED 4	–	–	0,073±0,012	–	0,069±0,012	0,069±0,011

Tabla A.59: Error medio absoluto para el regresor SVR rbf en el caso del experimento tipo 1 para cada combinación de LEDs 1.

MAE – Coord.	LEDS			
	1	2	3	4
Coord. X (m)	0,348±0,037	0,426±0,058	0,349±0,042	0,438±0,056
Coord. Y (m)	0,233±0,031	0,34±0,056	0,258±0,038	0,375±0,063

Tabla A.60: Error medio absoluto para el regresor SVR rbf en el caso del experimento tipo 1 para cada combinación de LEDs 2.

MAE – Coord.	LEDS					
	12	13	14	23	24	34
Coord. X (m)	0,107±0,021	0,363±0,047	0,12±0,022	0,13±0,031	0,415±0,069	0,104±0,02
Coord. Y (m)	0,076±0,014	0,248±0,035	0,099±0,018	0,098±0,017	0,356±0,058	0,099±0,025

Tabla A.61: Error medio absoluto para el regresor SVR rbf en el caso del experimento tipo 2 para cada combinación de LEDs 1.

MAE – Dist.(m)	LEDs			
	1	2	3	4
LED 1	0,071±0,01	–	–	–
LED 2	–	0,056±0,006	–	–
LED 3	–	–	0,062±0,01	–
LED 4	–	–	–	0,05±0,006

Tabla A.62: Error medio absoluto para el regresor SVR rbf en el caso del experimento tipo 2 para cada combinación de LEDs 2.

MAE – Dist.(m)	LEDs					
	12	13	14	23	24	34
LED 1	0,07±0,012	0,073±0,016	0,081±0,016	–	–	–
LED 2	0,065±0,008	–	–	0,068±0,009	0,068±0,013	–
LED 3	–	0,071±0,01	–	0,066±0,011	–	0,073±0,01
LED 4	–	–	0,059±0,008	–	0,055±0,009	0,056±0,008

Tabla A.63: Error relativo medio para el regresor SVR rbf en el caso del experimento tipo 1 para cada combinación de LEDs 1.

MAER – Coord.	LEDs			
	1	2	3	4
Coord. X (m)	0,25±0,042	0,377±0,099	0,306±0,078	0,316±0,064
Coord. Y (m)	0,195±0,042	0,328±0,098	0,175±0,027	0,237±0,035

Tabla A.64: Error relativo medio para el regresor SVR rbf en el caso del experimento tipo 1 para cada combinación de LEDs 2.

MAER – Coord.	LEDs					
	12	13	14	23	24	34
Coord. X (m)	0,098±0,032	0,298±0,08	0,071±0,013	0,16±0,068	0,354±0,105	0,072±0,016
Coord. Y (m)	0,072±0,025	0,195±0,039	0,071±0,015	0,086±0,028	0,292±0,099	0,061±0,014

Tabla A.65: Error relativo medio para el regresor SVR rbf en el caso del experimento tipo 2 para cada combinación de LEDs 1.

MAER – Dist.(m)	LEDs			
	1	2	3	4
LED 1	0,061±0,007	–	–	–
LED 2	–	0,047±0,006	–	–
LED 3	–	–	0,072±0,016	–
LED 4	–	–	–	0,05±0,018

Tabla A.66: Error relativo medio para el regresor SVR rbf en el caso del experimento tipo 2 para cada combinación de LEDs 2.

MAER – Dist.(m)	LEDs					
	12	13	14	23	24	34
LED 1	0,057±0,007	0,064±0,02	0,067±0,02	–	–	–
LED 2	0,051±0,007	–	–	0,054±0,007	0,054±0,009	–
LED 3	–	0,073±0,014	–	0,066±0,02	–	0,078±0,015
LED 4	–	–	0,055±0,025	–	0,054±0,022	0,054±0,022

Tabla A.67: Error medio absoluto para el regresor SVR rbf en el caso del experimento tipo 1 para cada combinación de LEDs 3.

MAE – Coord.	LEDS				
	123	124	134	234	1234
Coord. X (m)	0,107±0,022	0,08±0,013	0,093±0,017	0,075±0,015	0,074±0,011
Coord. Y (m)	0,083±0,014	0,077±0,017	0,096±0,02	0,088±0,017	0,084±0,015

Tabla A.68: Error medio absoluto para el regresor SVR rbf en el caso del experimento tipo 2 para cada combinación de LEDs 3.

MAE – Dist.(m)	LEDS				
	123	124	134	234	1234
LED 1	0,07±0,014	0,059±0,011	0,079±0,018	–	0,068±0,01
LED 2	0,068±0,009	0,074±0,012	–	0,074±0,014	0,074±0,012
LED 3	0,068±0,012	–	0,074±0,011	0,069±0,008	0,065±0,012
LED 4	–	0,063±0,009	0,062±0,007	0,066±0,008	0,062±0,011

Tabla A.69: Error medio absoluto para el regresor SVR rbf en el caso del experimento tipo 2 para cada combinación de LEDs.

MAE – Dist.(m)(Coord.)	LEDS				
	123	124	134	234	1234
Coord. X (m)	0,101±0,021	0,07±0,01	0,085±0,016	0,069±0,009	0,062±0,007
Coord. Y (m)	0,071±0,014	0,071±0,014	0,092±0,02	0,072±0,016	0,07±0,015

Tabla A.70: Error relativo medio para el regresor SVR rbf en el caso del experimento tipo 1 para cada combinación de LEDs 3.

MAER – Coord.	LEDS				
	123	124	134	234	1234
Coord. X (m)	0,117±0,049	0,056±0,013	0,063±0,015	0,069±0,032	0,057±0,019
Coord. Y (m)	0,073±0,024	0,065±0,022	0,066±0,014	0,066±0,017	0,065±0,02

Tabla A.71: Error relativo medio para el regresor SVR rbf en el caso del experimento tipo 2 para cada combinación de LEDs 3.

MAER – Dist.(m)	LEDS				
	123	124	134	234	1234
LED 1	0,059±0,014	0,055±0,017	0,062±0,022	–	0,068±0,014
LED 2	0,056±0,009	0,058±0,01	–	0,057±0,01	0,075±0,014
LED 3	0,068±0,016	–	0,073±0,015	0,072±0,016	0,064±0,024
LED 4	–	0,062±0,026	0,063±0,023	0,064±0,022	0,063±0,028

Tabla A.72: Error relativo medio para el regresor SVR rbf en el caso del experimento tipo 2 para cada combinación de LEDs.

MAER – Dist.(m) (Coord.)	LEDS				
	123	124	134	234	1234
Coord. X (m)	$0,104 \pm 0,045$	$0,048 \pm 0,011$	$0,055 \pm 0,011$	$0,05 \pm 0,011$	$0,044 \pm 0,01$
Coord. Y (m)	$0,066 \pm 0,026$	$0,06 \pm 0,02$	$0,062 \pm 0,011$	$0,056 \pm 0,014$	$0,057 \pm 0,019$

### A.2.8. Gráficas del Regresor SVR rbf

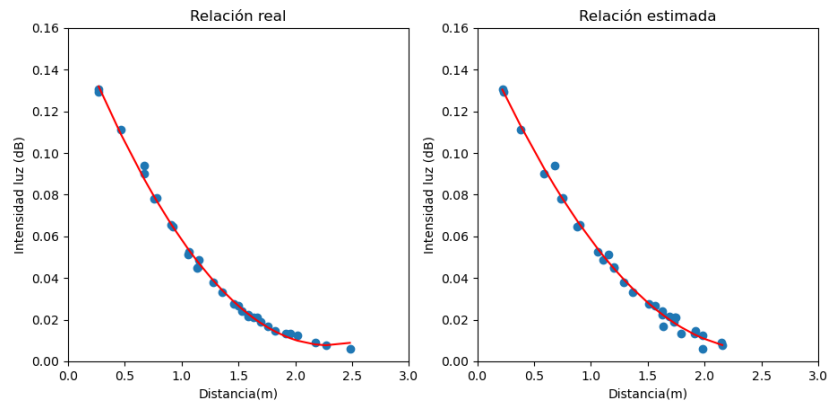


Figura A.35: Relación Intensidad-Distancia de los datos reales y los estimados para el LED 2.

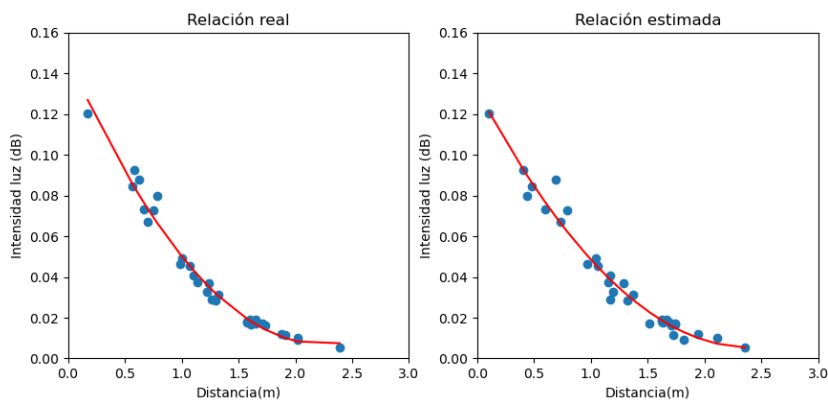


Figura A.36: Relación Intensidad-Distancia de los datos reales y los estimados para el LED 3.



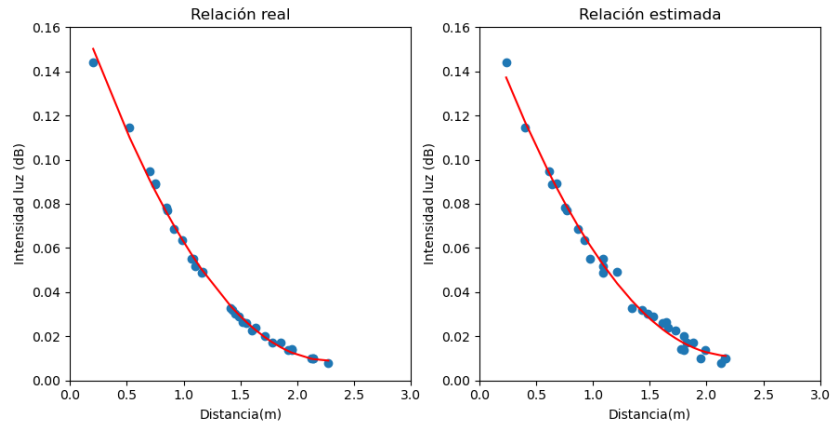


Figura A.37: Relación Intensidad-Distancia de los datos reales y los estimados para el LED 4.

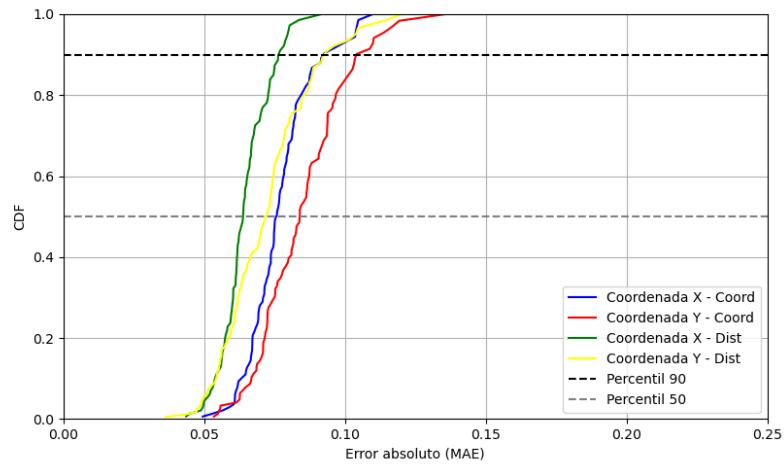


Figura A.38: CDF del MAE para el experimento tipo 2.

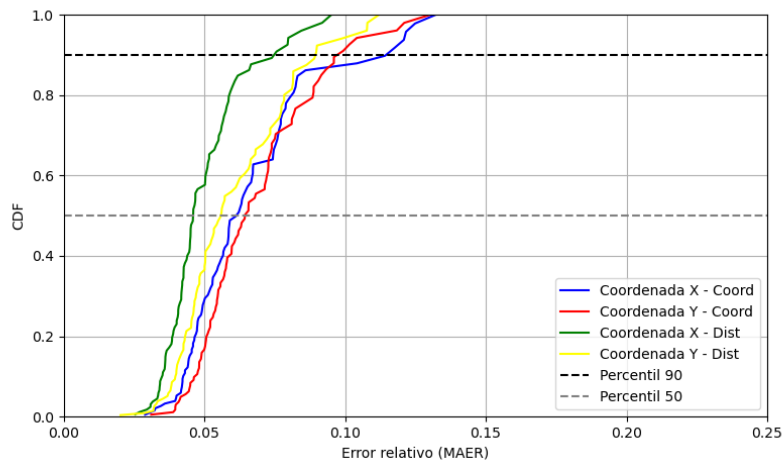


Figura A.39: CDF del MAER para el experimento tipo 1.

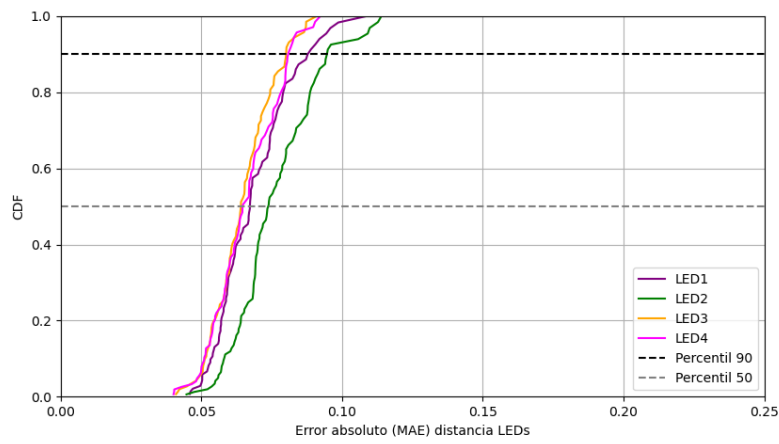


Figura A.40: CDF del MAE para el experimento tipo 2.

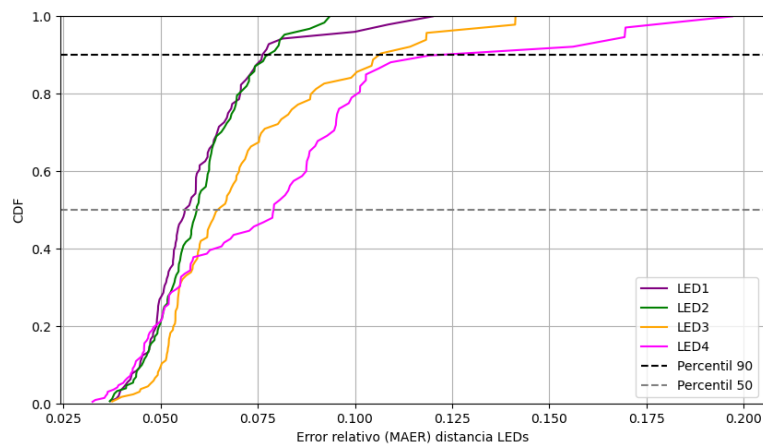


Figura A.41: CDF del MAER para el experimento tipo 2.

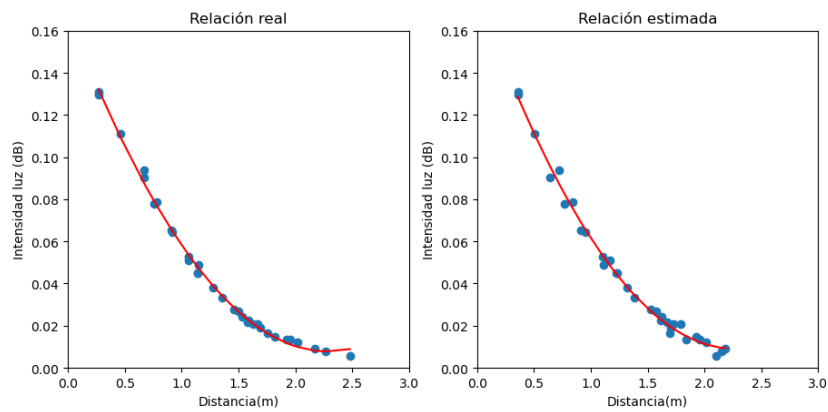


Figura A.42: Relación Intensidad-Distancia de los datos reales y los estimados para el LED 2 para el experimento tipo 2.

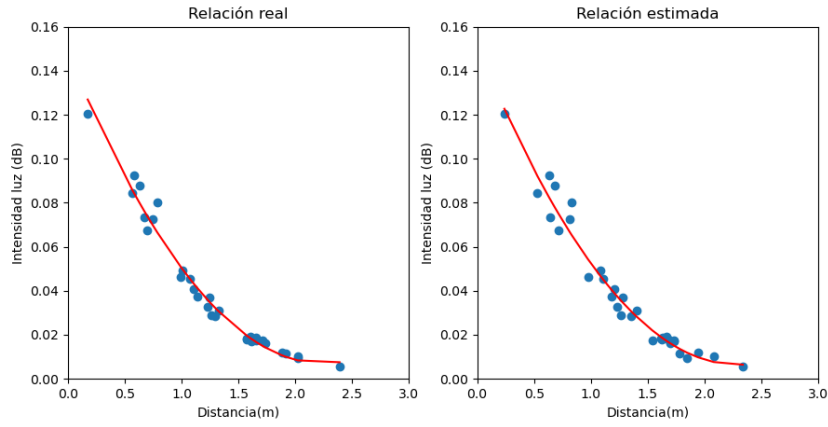


Figura A.43: Relación Intensidad-Distancia de los datos reales y los estimados para el LED 3 para el experimento tipo 2.

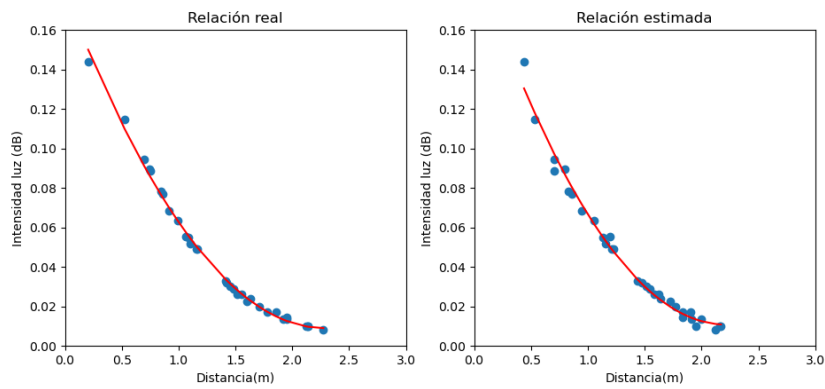
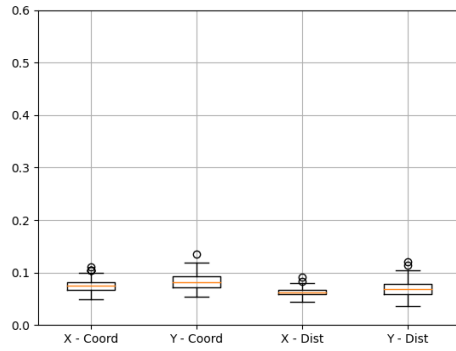
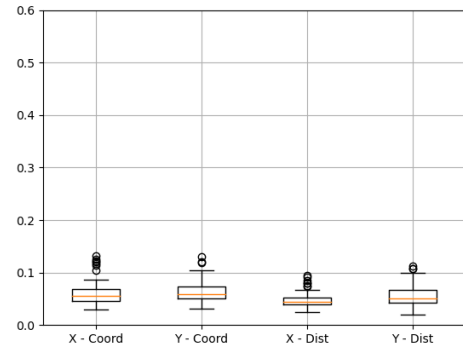


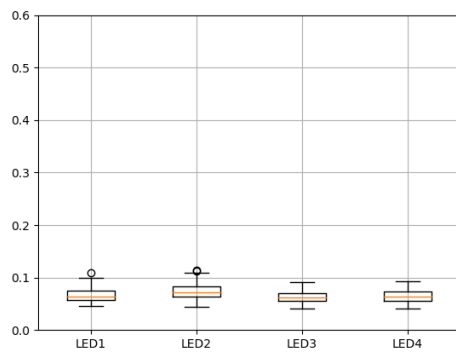
Figura A.44: Relación Intensidad-Distancia de los datos reales y los estimados para el LED 4 para el experimento tipo 2.



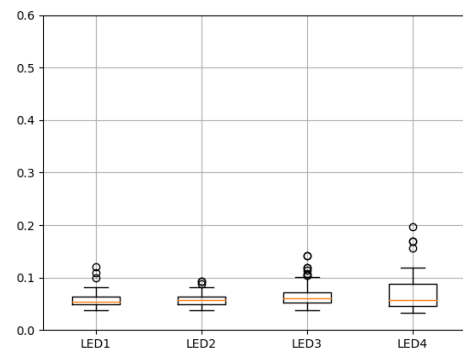
(a) Box-plot del MAE para el experimento tipo 1.



(b) Box-plot del MAER para el experimento tipo 1.



(c) Box-plot del MAE para el experimento tipo 2.



(d) Box-plot del MAER para el experimento tipo 2.

Figura A.45: Box-plot de los MAE y MAER para el regresor SVR RBF.

# Bibliografía

- [1] Anahid Basiri, Elena Simona Lohan, Terry Moore, Adam Winstanley, Pekka Peltola, Chris Hill, Pouria Amirian, and Pedro Figueiredo e Silva. Indoor location based services challenges, requirements and usability of current solutions. *Computer Science Review*, 24:1–12, 2017.
- [2] Martin Jacob, Sebastian Priebe, Robert Dickhoff, Thomas Kleine-Ostmann, Thorsten Schrader, and Thomas Kurner. Diffraction in mm and Sub-mm Wave Indoor Propagation Channels. *IEEE Transactions on Microwave Theory and Techniques*, 60(3):833–844, 2012.
- [3] Álvaro De-La-Llana-Calvo, José-Luis Lázaro-Galilea, Alfredo Gardel-Vicente, Ignacio Bravo-Muñoz, Aitor Alcázar-Fernández, Andreea Iamitchi, Rubén Gil-Vera, and David Hernández-Puerta. Evaluation of the accuracy and precision of orientation determination in an indoor positioning system using infrastructure illumination and PSD sensors. In *2021 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, pages 1–6, 2021.
- [4] M.M. Aman, G.B. Jasmon, H. Mokhlis, and A.H.A. Bakar. Analysis of the performance of domestic lighting lamps. *Energy Policy*, 52:482–500, 2013. Special Section: Transition Pathways to a Low Carbon Economy.
- [5] Yu-Cheng Chuang, Zhi-Qing Li, Chin-Wei Hsu, Yang Liu, and Chi-Wai Chow. Visible light communication and positioning using positioning cells and machine learning algorithms. *Opt. Express*, 27(11):16377–16383, May 2019.
- [6] Y. Xin and V. Ojanen. The impact of digitalization on product lifecycle management: How to deal with it? In *2017 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, pages 1098–1102, 2017.
- [7] Lalit Chettri and Rabindranath Bera. A Comprehensive Survey on Internet of Things (IoT) Toward 5G Wireless Systems. *IEEE Internet of Things Journal*, 7(1):16–32, 2020.

- [8] Lu Zhang, Jianjun Tan, Dan Han, and Hao Zhu. From machine learning to deep learning: progress in machine intelligence for rational drug discovery. *Drug discovery today*, 22(11):1680–1685, 2017.
- [9] Ahmed Amer and Phillip Peralez. Affordable altered perspectives: Making augmented and virtual reality technology accessible. In *IEEE Global Humanitarian Technology Conference (GHTC 2014)*, pages 603–608, 2014.
- [10] Weipeng Guan, Shihuan Chen, Shangsheng Wen, Zequn Tan, Hongzhan Song, and Wenyuan Hou. High-Accuracy Robot Indoor Localization Scheme Based on Robot Operating System Using Visible Light Positioning. *IEEE Photonics Journal*, 12(2):1–16, 2020.
- [11] Yuan Zhuang, Qin Wang, Min Shi, Pan Cao, Longning Qi, and Jun Yang. Low-Power Centimeter-Level Localization for Indoor Mobile Robots Based on Ensemble Kalman Smoother Using Received Signal Strength. *IEEE Internet of Things Journal*, 6(4):6513–6522, 2019.
- [12] R. Perez-Jimenez, J. Rabadan, V. Guerra, L. Aguiar, and J. Rufo. Fundamentals of Indoor Vlp: Providing Autonomous Mobility for Visually Impaired People. In *2017 International Conference and Workshop on Bioinspired Intelligence (IWObI)*, pages 1–6, 2017.
- [13] Daniel Konings, Nathaniel Faulkner, Fakhrul Alam, Edmund M.-K. Lai, and Serge Demidenko. FieldLight: Device-Free Indoor Human Localization Using Passive Visible Light Positioning and Artificial Potential Fields. *IEEE Sensors Journal*, 20(2):1054–1066, 2020.
- [14] Xu Yang, Xin Yang, and Hui-Lin Xiong. Multi-view face detection with the multi-resolution MPP classifiers. In *2009 International Conference on Machine Learning and Cybernetics*, volume 3, pages 1770–1775, 2009.
- [15] Srividya Kotagiri, Attada Venkataramana, and Gogula Kiran. Blind Aid: State of the art for Scene Text Detector and Text to Speech. In *2022 International Conference on Advanced Computing Technologies and Applications (ICACTA)*, pages 1–5, 2022.
- [16] Ittikon Thammachantuek, Somkiat Kosolsomnbat, and Mahasak Ketcham. Comparison of Machine Learning Algorithm’s Performance Based on Decision making in Autonomous Car. In *2018 International Joint Symposium on Artificial Intelligence and Natural Language Processing (iSAI-NLP)*, pages 1–6, 2018.
- [17] S Jayabharathi and V Ilango. A Brief Revolution of Evolution and Resurgence on Machine Learning. In *2021 Asian Conference on Innovation in Technology (ASIAN-CON)*, pages 1–5, 2021.

- 
- [18] George Strawn. Alan Turing. *IT Professional*, 16(1):5–7, 2014.
- [19] E.A. Weiss. Biographies: Eloge: Arthur Lee Samuel (1901-90). *IEEE Annals of the History of Computing*, 14(3):55–69, 1992.
- [20] Bernard Widrow, Reiner Hartenstein, and Robert Hecht Nielsen. 1917 Karl Steinbuch 2005.
- [21] Frank Rosenblatt. Perceptron Simulation Experiments. *Proceedings of the IRE*, 48(3):301–309, 1960.
- [22] Andreja Kitanov, Vedran Tubin, and Ivan Petrovic. Extending functionality of RF Ultrasound positioning system with dead-reckoning to accurately determine mobile robot’s orientation. In *2009 IEEE Control Applications, (CCA) & Intelligent Control, (ISIC)*, pages 1152–1157, 2009.
- [23] Jasurbek Khodjaev, Anvar Narzullaev, Yongwan Park, Wooyoung Jung, Jonghun Lee, and Sangdong Kim. Performance Improvement of Asynchronous UWB Position Location Algorithm Using Multiple Pulse Transmission. In *2007 4th Workshop on Positioning, Navigation and Communication*, pages 167–170, 2007.
- [24] Zhiwei Liu, Jiansheng Li, Ancheng Wang, Xiangbo Cheng, and Anran Wang. Design and Implementation of UWB/MIMU Tightly-coupled System for Indoor Positioning. In *2018 Ubiquitous Positioning, Indoor Navigation and Location-Based Services (UPINLBS)*, pages 1–7, 2018.
- [25] Youngsu Cho, Myungin Ji, Yangkoo Lee, and Sangjoon Park. WiFi AP position estimation using contribution from heterogeneous mobile devices. In *Proceedings of the 2012 IEEE/ION Position, Location and Navigation Symposium*, pages 562–567, 2012.
- [26] Ling Pei, Ruizhi Chen, Jingbin Liu, Tomi Tenhunen, Heidi Kuusniemi, and Yuwei Chen. An Inquiry-based Bluetooth indoor positioning approach for the Finnish pavilion at Shanghai World Expo 2010. In *IEEE/ION Position, Location and Navigation Symposium*, pages 1002–1009, 2010.
- [27] Huaping Li, Hongbin Huang, Yongze Xu, Zhanhang Wei, Sichen Yuan, Puxi Lin, Hao Wu, Wen Lei, Junbin Fang, and Zhe Chen. A Fast and High-Accuracy Real-Time Visible Light Positioning System Based on Single LED Lamp With a Beacon. *IEEE Photonics Journal*, 12(6):1–12, 2020.
- [28] Heidi Steendam, Thomas Qian Wang, and Jean Armstrong. Theoretical Lower Bound for Indoor Visible Light Positioning Using Received Signal Strength Measurements

- and an Aperture-Based Receiver. *Journal of Lightwave Technology*, 35(2):309–319, 2017.
- [29] Adli Hasan Abu Bakar, Tyrel Glass, Hing Yan Tee, Fakhurul Alam, and Mathew Legg. Accurate Visible Light Positioning Using Multiple-Photodiode Receiver and Machine Learning. *IEEE Transactions on Instrumentation and Measurement*, 70:1–12, 2021.
- [30] Yu-Chun Wu, Chi-Wai Chow, Yang Liu, Yun-Shen Lin, Chong-You Hong, Dong-Chang Lin, Shao-Hua Song, and Chien-Hung Yeh. Received-Signal-Strength (RSS) Based 3D Visible-Light-Positioning (VLP) System Using Kernel Ridge Regression Machine Learning Algorithm With Sigmoid Function Data Preprocessing Method. *IEEE Access*, 8:214269–214281, 2020.
- [31] Chong-You Hong, Yu-Chun Wu, Yang Liu, Chi-Wai Chow, Chien-Hung Yeh, Ke-Ling Hsu, Dong-Chang Lin, Xin-Lan Liao, Kun-Hsien Lin, and Yi-Yuan Chen. Angle-of-Arrival (AOA) Visible Light Positioning (VLP) System Using Solar Cells With Third-Order Regression and Ridge Regression Algorithms. *IEEE Photonics Journal*, 12(3):1–5, 2020.
- [32] Shao-Hua Song, Dong-Chang Lin, Yang Liu, Chi-Wai Chow, Yun-Han Chang, Kun-Hsien Lin, Yi-Chang Wang, and Yi-Yuan Chen. Employing DIALux to relieve machine-learning training data collection when designing indoor positioning systems. *Opt. Express*, 29(11):16887–16892, May 2021.
- [33] Tao Yuan, Yiqin Xu, Yong Wang, Peng Han, and Junfang Chen. A Tilt Receiver Correction Method for Visible Light Positioning Using Machine Learning Method. *IEEE Photonics Journal*, 10(6):1–12, 2018.
- [34] Canyu Xie, Weipeng Guan, Yuxiang Wu, Liangtao Fang, and Ye Cai. The LED-ID Detection and Recognition Method Based on Visible Light Positioning Using Proximity Method. *IEEE Photonics Journal*, 10(2):1–16, 2018.
- [35] Yu-Chun Wu, Ke-Ling Hsu, Yang Liu, Chong-You Hong, Chi-Wai Chow, Chien-Hung Yeh, Xin-Lan Liao, Kun-Hsien Lin, and Yi-Yuan Chen. Using Linear Interpolation to Reduce the Training Samples for Regression Based Visible Light Positioning System. *IEEE Photonics Journal*, 12(2):1–5, 2020.
- [36] Seyedeh Leili Mirtaheri and Reza Shahbazian. *Machine learning: theory to applications*. CRC Press, 2022.
- [37] Lina Zhou, Shimei Pan, Jianwu Wang, and Athanasios V Vasilakos. Machine learning on big data: Opportunities and challenges. *Neurocomputing*, 237:350–361, 2017.



- [38] Han Sun, Henry V Burton, and Honglan Huang. Machine learning applications for building structural design and performance assessment: State-of-the-art review. *Journal of Building Engineering*, 33:101816, 2021.
- [39] Q. Wang. *Learning to Understand Remote Sensing Images*. Mdpi AG, 2019.
- [40] Md Zahangir Alom, Tarek M Taha, Chris Yakopcic, Stefan Westberg, Paheding Sidi-ke, Mst Shamima Nasrin, Mahmudul Hasan, Brian C Van Essen, Abdul AS Awwal, and Vijayan K Asari. A state-of-the-art survey on deep learning theory and architectures. *electronics*, 8(3):292, 2019.
- [41] M.C. Thrun. *Projection-Based Clustering through Self-Organization and Swarm Intelligence: Combining Cluster Analysis with the Visualization of High-Dimensional Data*. Springer Fachmedien Wiesbaden, 2018.
- [42] M. Awad and R. Khanna. *Efficient Learning Machines: Theories, Concepts, and Applications for Engineers and System Designers*. Books for professionals by professionals. Apress, 2015.
- [43] Daniel Berrar. Cross-validation. *Encyclopedia of Bioinformatics and Computational Biology*, 1(April):542–545, 2019.
- [44] Wikipedia. Validación cruzada — wikipedia, la enciclopedia libre, 2023. [Internet; descargado 16-abril-2023].
- [45] Tzu-Tsung Wong. Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation. *Pattern recognition*, 48(9):2839–2846, 2015.
- [46] K.R. Watts, Geological Survey (U.S.), and United States. Bureau of Reclamation. *Regression Models of Monthly Water-level Change in and Near the Closed Basin Division of the San Luis Valley, South-central Colorado*. Water-resources investigations report. U.S. Department of the Interior, U.S. Geological Survey, 1995.
- [47] J. Nzabanita. *Bilinear and Trilinear Regression Models with Structured Covariance Matrices*. Linköping Studies in Science and Technology. Dissertations. Linköping University Electronic Press, 2015.
- [48] D.O. Loftsgaarden, P.L. Andrews, and Utah) Intermountain Research Station (Ogden. *Constructing and Testing Logistic Regression Models for Binary Data: Applications to the National Fire Danger Rating System*. General technical report INT. U.S. Department of Agriculture, Forest Service, Intermountain Research Station, 1992.
- [49] D.J. Lystrom, Ga.) Environmental Research Laboratory (Athens, and Geological Survey (U.S.). *Multiple Regression Modeling Approach for Regional Water Quality Management*. Number v. 1 in Multiple Regression Modeling Approach for Regional

- Water Quality Management. Environmental Protection Agency, Office of Research and Development, Environmental Research Laboratory, 1978.
- [50] Kavitha S, Varuna S, and Ramya R. A comparative analysis on linear regression and support vector regression. In *2016 Online International Conference on Green Engineering and Technologies (IC-GET)*, pages 1–5, 2016.
- [51] B. Üstün, W.J. Melssen, and L.M.C. Buydens. Facilitating the application of Support Vector Regression by using a universal Pearson VII function based kernel. *Chemo-metrics and Intelligent Laboratory Systems*, 81(1):29–40, 2006.
- [52] Mufni Alida and Metty Mustikasari. Rupiah Exchange Prediction of US Dollar Using Linear, Polynomial, and Radial Basis Function Kernel in Support Vector Regression. *Jurnal Online Informatika*, 5(1):53–60, 2020.
- [53] Rezzy Eko Caraka, Rung Ching Chen, Toni Toharudin, Muhammad Tahmid, Bens Pardamean, and Richard Mahendra Putra. Evaluation performance of SVR genetic algorithm and hybrid PSO in rainfall forecasting. *ICIC Express Lett Part B Appl*, 11(7):631–639, 2020.
- [54] H. Tabari. *Statistical Analysis and Stochastic Modelling of Hydrological Extremes*. MDPI AG, 2019.
- [55] Jiandong Huang, Yuantian Sun, and Junfei Zhang. Reduction of computational error by optimizing SVR kernel coefficients to simulate concrete compressive strength through the use of a human learning optimization algorithm. *Engineering with Computers*, pages 1–18, 2021.
- [56] Kui Fu Chen. Estimating Parameters of a Sine Wave by Separable Nonlinear Least Squares Fitting. *IEEE Transactions on Instrumentation and Measurement*, 59(12):3214–3217, 2010.
- [57] M. Merriman. *A Text Book on the Method of Least Squares*. J. Wiley, 1884.
- [58] Willem Raes, Jorik De Bruycker, Nicolas Knudde, Tom Dhaene, and Nobby Stevens. Machine Learning for RSS-Based Visible Light Positioning. 2020.
- [59] Nicolas Knudde, Willem Raes, Jorik De Bruycker, Tom Dhaene, and Nobby Stevens. Data-efficient gaussian process regression for accurate visible light positioning. *IEEE Communications Letters*, 24(8):1705–1709, 2020.