



UNIVERSIDAD DE BURGOS
ESCUELA POLITÉCNICA SUPERIOR
Grado en Ingeniería Informática



**TFG del Grado en Ingeniería
Informática**

**Creación de demos técnica y
documentación asociada, para
la detección/clasificación de
materiales, utilizando un
sensor radar a 60GHz**



Presentado por Martín Encabo Contreras
en Universidad de Burgos — 13 de enero
de 2022

Tutores:

Pedro Latorre Carmona
José Francisco Díez Pastor



UNIVERSIDAD DE BURGOS
ESCUELA POLITÉCNICA SUPERIOR
Grado en Ingeniería Informática



D. nombre tutor, profesor del departamento de nombre departamento, área de nombre área.

Expone:

Que el alumno D. Martín Encabo Contreras, con DNI 72897369L, ha realizado el Trabajo final de Grado en Ingeniería Informática titulado título de TFG.

Y que dicho trabajo ha sido realizado por el alumno bajo la dirección del que suscribe, en virtud de lo cual se autoriza su presentación y defensa.

En Burgos, 13 de enero de 2022

Vº. Bº. del Tutor:

Vº. Bº. del co-tutor:

D. Pedro Latorre Carmona

D. José Francisco Díez Pastor

Resumen

Con el desarrollo de la tecnología se han obtenido diferentes avances tecnológicos los cuales han beneficiado las actividades de la humanidad es por ello que en el siguiente proyecto se propone el desarrollo de una aplicación la cual opera mediante código Python el cual permite que por medio de la obtención de lecturas de diferentes de materiales sea posible el clasificarlos e indicarlos, definiendo el tipo de material, todo ello implementando el uso de un radar a una frecuencia de 60 GHz, con el cual es posible indicar cual es la probabilidad este presente el material en la lectura.

Dado que es necesario realizar un análisis de los datos obtenidos por medio de radar, se propone implementar técnicas de aprendizaje automático para poder reconocer los diferentes materiales que se estudiaran, es por ello por lo que se propone el algoritmo clasificador Random Forest en conjunto con la validación cruzada.

Para poder recopilar la información necesaria, será necesario hacer uso de la librería del radar que es proporcionada por el fabricante, esto con el objetivo de poder conseguir una serie de lecturas, las cuales deberán brindar la suficiente información para poder clasificar, indicar e identificar todos los objetos sometidos al estudio, se propone un análisis con un registro de treinta materiales considerando diez lecturas generadas de cada material.

Descriptores

Aprendizaje Automático, Clasificador, Random Forest, Reconocimiento, Python, Sensor, Radar, Acconeer.

Abstract

With the development of technology, different technological advances have been obtained which have benefited the activities of humanity, which is why in the following project the development of an application is proposed which operates through Python code which allows through the Obtaining readings of different materials, it is possible to classify and indicate them, defining the type of material, all this by implementing the use of a radar at a frequency of 60 GHz, with which it is possible to indicate what is the probability that the material is present in the reading.

Since it is necessary to carry out an analysis of the data obtained by means of radar, it is proposed to implement machine learning techniques to be able to recognize the different materials that will be studied, that is why the Random Forest classifier algorithm is proposed in conjunction with the cross validation.

In order to collect the necessary information, it will be necessary to make use of the radar library that is provided by the manufacturer, this in order to be able to obtain a series of readings, which must provide enough information to be able to classify, indicate and identify all the objects submitted to the study, an analysis is proposed with a record of thirty materials considering ten generated readings of each material.

Keywords

Machine Learning, Classifier, Random Forest, Recognition, Python, Sensor, Radar, Acconeer.

Índice general

| | |
|--|-----|
| Índice general | iii |
| Índice de figuras | v |
| Índice de tablas | vi |
| Introducción | 1 |
| Objetivos del proyecto | 3 |
| 2.1. Objetivos generales | 3 |
| 2.2. Objetivos técnicos | 4 |
| Conceptos teóricos | 7 |
| 3.1. Radar | 7 |
| 3.2. Teledetección | 10 |
| 3.3. Radar Acconeer | 16 |
| 3.4. Servicio In-phase and Quadrature | 22 |
| 3.5. Aprendizaje automático | 25 |
| 3.6. Algoritmos machine learning | 35 |
| 3.7. Características de los materiales | 40 |
| Técnicas y herramientas | 43 |
| 4.1. Técnicas y metodologías | 43 |
| 4.2. Lenguajes y bibliotecas | 44 |
| 4.3. Herramientas de desarrollo | 46 |
| 4.4. Herramientas de documentación | 46 |
| Aspectos relevantes del desarrollo del proyecto | 47 |

| | |
|---|-----------|
| 5.1. Propuesta del Proyecto | 47 |
| 5.2. Metodologías aplicadas | 48 |
| 5.3. Formación | 48 |
| 5.4. Montaje del radar | 48 |
| 5.5. Lectura de los objetos | 50 |
| 5.6. Modelo clasificador | 52 |
| Trabajos relacionados | 53 |
| 6.1. RadarCat | 53 |
| Conclusiones y Líneas de trabajo futuras | 55 |
| 7.1. Conclusiones | 55 |
| 7.2. Líneas de trabajo futuras | 55 |
| Bibliografía | 57 |

Índice de figuras

| | |
|---|----|
| 3.1. Radar A111. | 17 |
| 3.2. Diagrama de bloques del sensor A111. | 20 |
| 3.3. Random Forest | 37 |
| 3.4. K-nearest neighbors | 39 |
| 3.5. Herramienta Acconeer | 40 |
| 5.1. Regresión Logística | 52 |
| 6.1. RadarCat. | 54 |

Índice de tablas

| | |
|---|----|
| 3.1. Servicios del radar A111 | 22 |
|---|----|

Introducción

La tecnología de radar existe desde la década de 1930 de la mano de Watson-Watt, el término radar proviene del acrónimo inglés RAdio Detection And Ranging, el uso típico de los radares de radiofrecuencia se basa en medir distancias, altitudes, direcciones y velocidades de objetos, además de tener un uso muy reconocido el cual es el mapa de navegación de los barcos.

Durante los últimos años han surgido nuevas áreas de aplicación las cuales plantean diferentes desafíos para la humanidad, dado que una aplicación es monitorizar los signos vitales, reconocimiento de gestos, entre otros, dado que los radares últimamente se han vuelto más baratos, en gran parte debido a su adopción en la industria automotriz, lo que convierte a estos dispositivos en una opción atractiva en una amplia gama de aplicaciones de bajo costo.

El presente trabajo está basado en documentar, implementar, tratando de demostrar que mediante el uso de un radar de 60GHz fabricado por Acconeer es posible el reconocimiento de diferentes materiales, por lo cual se plantea el implementar un procedimiento el cual ayude a la obtención de las diferentes características, utilizando algoritmos de aprendizaje automático, dado que se busca realizará un registro de tres tipos de materiales (cartón, plástico y cristal) para crear un modelo de Inteligencia Artificial para conseguir un correcto funcionamiento del radar.

Objetivos del proyecto

El objetivo de desarrollar este proyecto es el poder implementar la tecnología del radar para poder verificar su funcionamiento y operación del sensor radar por medio de una herramienta de machine learning capaz de establecer una comparación entre los diferentes tipos de materiales que se estudiaran.

2.1. Objetivos generales

Los objetivos generales del proyecto que caracterizan la propuesta que se presenta a continuación se puede describir en los siguientes puntos:

- El objetivo principal del proyecto es poder documentar los diferentes elementos que componen el ensamblar y poner en funcionamiento un radar de 60GHz fabricado por Acconeer el cual es capaz de identificar, mediante un clasificador, el tipo de material al que pertenece un objeto.
- Crear bajo un algoritmo de machine learning un procedimiento capaz de extraer determinadas características de los materiales, con el empleo del aprendizaje automático.
- Poder generar un gráfico de probabilidades el cual sea capaz de mostrar la clasificación de los diferentes materiales empleados.
- Se establecerá una interfaz capaz de poder ser implementada en el sistema operativo Windows, buscando que la aplicación pueda ser implementada en cualquier dispositivo que cuente con este sistema.

2.2. Objetivos técnicos

Las diferentes ramas empleadas para desarrollar el proyecto hacen que sea un proyecto multidisciplinario, ya que está integrado por ramas de la electrónica, sistemas computacionales, física, mecánica, mecatrónica, por lo cual se presentan los siguientes objetivos técnicos:

- Conocer e identificar las diferentes tecnologías empleadas en la telede-tección, describiendo que es un radar, cuál es su operación, cuales son sus limitaciones, logrando aterrizar la idea principal del proyecto del empleo de un radar para el estudio.
- Conocer los datos técnicos del radar acconeer el cual será implementado para establecer cuáles serán sus requerimientos en cuanto a alimentación eléctrica, conexión, interacción con los demás componentes que integran el proyecto
- Logra establecer las diferencias entre los diferentes aprendizajes auto-máticos existentes para poder determinar cuál es la mejor opción de acuerdo con la información que se pretende obtener de los diferentes materiales.
- Establecer un algoritmo machine learning el cual pueda simplificar la operación de identificación de los diferentes materiales haciéndolo de una forma automática.
- Establecer las diferentes características a analizad de los materiales buscando poder diferenciarlos con el empleo del sensor radar acconeer, para establecer la base de datos generada.
- Establecer cuales técnicas y metodologías se deben implementar para poder establecer los tiempos y movimientos para realizar las revisiones y así poder obtener los resultados deseados en el proyecto
- Determinar que lenguajes serán los óptimos para ser empleados de acuerdo con los componentes y los conocimientos obtenidos durante el análisis de la literatura que integra este proyecto.
- Se establecerá cual herramienta de documentación será empleada para la elaboración de documentos que integran el proyecto pudiendo editar, gestionar y difundir la información de una forma segura.

- Establecer las bases para futuras investigaciones las cuales deseen implementar este tipo de tecnología en sus proyectos, sirviendo como referencia y apoyo a los diferentes desarrolladores.

Conceptos teóricos

Para poder construir un modelo capaz de capturar con precisión las características distintivas en cuanto a las superficies objetivo, es necesario comprender el origen y la estructura de la señal recibida. En este apartado se introducen algunos conceptos fundamentales del sistema de radar.

3.1. Radar

El radar o mejor conocido por su acrónimo del inglés, radio detecting and ranging, se describe como la detección y localización por radio, el cual es un sistema electrónico el cual emplea radiaciones electromagnéticas, las cuales son reflejadas por un objeto para así poder determinar características del objeto como son su localización, la distancia a la que se encuentra y su velocidad, en dado caso que se encuentre en desplazamiento [1].

Durante el desarrollo tecnológico basado en la postulación de las teorías generadas sobre las ondas electromagnéticas se han desarrollado múltiples avances tecnológicos los cuales han hecho posible que en la actualidad se puedan emplear los radares en diferentes y diversas áreas con seguridad y eficiencia.

Los constantes avances tecnológicos en el desarrollo de materiales y su implementación en los circuitos integrados, junto con el desarrollo de la informática, son factores importantes los cuales han convertido a los sistemas de radares en instrumentos de bajo tamaño con una precisión y calidad.

Principio de funcionamiento del Radar

El radar es un sistema de teledetección activa la cual emite ondas de microondas, donde el valor de las frecuencias en donde están comprendidos los cuales tienen diferentes valores los cuales van entre 1 GHz y 100 GHz, estos valores también se pueden considerar como longitudes de onda, las cuales tienen valores de entre 3 mm y 30 cm [2].

Las ondas de microondas tienen diferentes ventajas, entre las cuales se puede considerar que son capaces de atravesar los cúmulos de nubes, además de poder atravesar también la lluvia, otra ventaja es que permiten trabajar en condiciones de oscuridad debido a que no dependen de la iluminación solar o artificial, otra ventaja es que pueden penetrar con ciertas limitaciones las estructuras vegetales, la nieve y el suelo, además de poder tener la ventaja de proporcionar una alta sensibilidad para poder medir características como la distancia, la rugosidad, la humedad, entre otras características de los materiales.

Los sistemas radar de apertura sintética, también conocidos como SAR, cuentan con un sensor radar el cual es capaz de emitir un pulso de ondas de microondas dirigida hacia la superficie del terreno u objeto, el cual en cierto momento retorna en dirección al sistema radar de apertura sintética que se encuentra ubicado en un satélite, donde la información conseguida es medida por el sensor radar.

Los pulsos empleados por el sistema radar de apertura sintética los cuales típicamente poseen longitudes de onda las cuales pueden variar desde los 3 cm de la banda X, hasta los 24 cm de la banda L, por lo que cuanto mayor sea la longitud de onda, mayor será su capacidad de penetración en zonas vegetadas, más sin embargo habría que considerar que será peor su resolución espacial.

Los sistemas radar de apertura sintética pueden ser transportados mediante un satélite, un avión o permanecer en una plataforma terrestre, en donde se emiten y reciben la señal del pulso de microondas emitido por el sensor radar, el cual está generando imágenes complejas de alta resolución espacial de la superficie del terreno a las cuales se les conoce como imágenes SAR. Podrían mencionarse diferentes casos en los cuales se generan imágenes complejas de alta resolución espacial como por ejemplo el satélite ERS-1/2, junto con el satélite ENVISAT, los cuales pertenecen a la Agencia Espacial Europea en estos satélites existe un sensor el cual emite ondas en la banda C, esta onda tiene una longitud de onda de 6 cm, con la cual es posible generar

imágenes SAR, las cuales pueden llegar a tener una resolución espacial de 4 m x 20 m.

Otro caso es el que encontramos en el satélite TERRASAR-X donde las ondas de microondas generadas pertenecen a la banda X, pudiendo generar imágenes SAR de 1 m x 3 m de resolución, además de que las imágenes tipo SAR, adquiridas por diferentes satélites los cuales trabajan en la banda C, las imágenes pueden llegar a tener un tamaño de 100 km x 100 km con una excelente resolución espacial, a cuál puede ser de 4 m x 20 m para cada píxel.

En las imágenes tipo SAR se puede considerar que cada píxel es representado con un formato de número complejo el cual contiene un valor de amplitud, junto con otro valor de fase, donde el valor de la amplitud está directamente relacionado con el coeficiente de respuesta el cual se generó con la información de la superficie del terreno dirigido a la señal de respuesta, esta señal también es ubicada como dispersión, además de que fase también conocida como φ , puede contener información sobre la distancia, la cual es recorrida por la señal, que inicia desde el sensor hasta la superficie del terreno, donde son imprescindibles la longitud de onda electromagnética λ y la distancia desde el satélite a la superficie del terreno R .

En la obtención de imágenes SAR una de las partes más importantes del radar es el emisor, el cual es un instrumento que produce la radiación electromagnética, donde la toda la energía emitida debe concentrarse en otra parte del radar la esta está conformada por un único lóbulo también conocido como principal, aunque en realidad existen pequeños lóbulos secundarios a su alrededor [3].

Tipos de radar

En el mercado y de acuerdo a su forma de operar existen diferentes tipos de radares los cuales pueden operar de modo primario o de modo secundario, considerando que todos los radares tienen la misma base de operabilidad ya que todos los radares tienen en común la capacidad de poder transmitir una señal, pudiendo generar la evaluación mediante diferentes técnicas de procesamiento de las señales recibidas para así poder obtener el parámetro buscado [4].

Los sistemas de radar por lo general se encuentran seccionados en diferentes categorías de operación las cuales están basada en los diferentes métodos de transmisión de la señal, por lo general estos métodos corresponden a la transmisión de pulsos, la onda continua y la frecuencia modulada, en

donde el método de transmisión de pulsos es el método más común para implementarse en la transmisión de la energía del radar.

El método de onda continua se basa en el principio del El Efecto Doppler, para poder generar la detección de la presencia y velocidad de un objeto en movimiento en dirección al radar o en dirección opuesta al radar, debido a que el sistema no puede determinar el rango, tampoco logra diferencia entre objetos y obstáculos que se encuentran en la misma dirección, los cuales están viajando a la misma velocidad, este método en los radares por lo general es empleado en los sistemas de control de fuego o disparo en el campo militar, para poder rastrear objetos los cuales presentan movimientos rápidos en un rango cercano.

3.2. Teledetección

La teledetección, también conocida como detección a distancia o detección remota, es la técnica que permite recopilar información a distancia de los objetos sin que exista un contacto físico, para lo cual es necesario tener una interacción entre los objetos a estudiar y un sensor situado en una plataforma.

El radar es un sensor activo de microondas el cual se desplaza a bordo de una plataforma explorando la tierra, por lo general se encuentra emitiendo pulsos de energía con dirección a la superficie terrestre, pudiendo almacenar las señales obtenidas de retorno [5].

Debido a que tienen que trabajar con haces de energía emitidos artificialmente, los sistemas de radar permiten controlar las condiciones de la adquisición, basadas en la frecuencia, la polarización y la geometría de la observación, ya pueden recolectar datos en cualquier condición ya sea de día o de noche, el realizar este tipo de trabajo resulta muy ventajoso en las regiones polares, lugares donde los prolongados períodos de oscuridad dificultan la adquisición de imágenes convencionales.

Por lo general un sistema de teledetección por RADAR (Radio Detection And Ranging) transmite diferentes pulsos de microondas los cuales van barriendo la superficie terrestre, las cuales reciben la porción de energía que es reflejada o retrodispersada, retornando de vuelta hacia el sensor del radar.

Al momento de que el radar recibe la señal el sistema registra la intensidad de la señal de retorno conocida como radiación retrodispersada, donde el retardo en tiempo entre la transmisión y recepción de cada pulso de energía

es el tiempo que se relaciona con la distancia de los objetos observados. La distancia a la cual debe estar situado un sensor para poder ser considerado remoto puede variar ya sea desde pocos decímetros hasta miles de kilómetros, ya que la teledetección es un flujo de radiación emitido por los objetos o materiales hacia un radar o sensor.

La teledetección es un flujo de radiación emitido por los objetos o materiales hacia un radar o sensor. El origen del flujo puede venir de:

- Radiación solar reflejada por los objetos - Teledetección pasiva
- Radiación terrestre emitida por los objetos - Teledetección pasiva
- Radiación emitida por el sensor y reflejada por los objetos (radar) - Teledetección activa

El grupo de datos adquiridos por medio de los diferentes procedimientos empleados en la Teledetección son generados desde plataformas como aviones o transbordadores espaciales, los cuales contienen por lo general tres tipos de información, la primera es la información espacial la cual es capaz de representar la organización de los elementos que constituyen la imagen en el espacio físico, el segundo tipo de información está caracterizada por contener información espectral la cual es posible caracteriza, de tal forma que se puede conducir a la obtención del conocimiento de la naturaleza en la superficie terrestre y el tercer tipo de información es una información de tipo temporal la cual permite generar la detección de los cambios producidos en la superficie del planeta Tierra, después del transcurso de determinado tiempo [6].

Por lo general los datos que son adquiridos a través de la Teledetección son caracterizados por algunas de las siguientes propiedades de los materiales los cuales requieren de la capacidad para obtener la información en base a los aspectos del medio natural los cuales pasan desapercibidos a nuestros sentidos, como son las ondas de radar, el infrarrojo de los satélites LANDSAT, entre otros.

Por lo que se puede observar la experiencia natural que tiene el hombre la cual por lo general nula si se consideran los dominios espectrales, debido a que es imperceptible es por lo que se desarrollan visualizaciones las cuales tienen una determinada función y utilidad, las cuales están en conjunto con las fotografías aéreas, las cuales reciben el nombre de imágenes para así evitar generar confusión.

Otra propiedad de los datos adquiridos es la generación de este tipo de informaciones las cuales son registradas por los sensores, las cuales son capaces de medir la cantidad de energía la cual es reflejada o emitida por medio de los objetos que se presentan en la naturaleza los cuales componen el paisaje, debido a que esta información es representada de forma numérica es como se da la capacidad de ser tratados en elementos matemáticos.

Pero al ser su existencia de extrema abundancia se ha optado por aplicar el uso de enormes ordenadores los cuales son capaces de aplicar métodos matemáticos que tienen la capacidad de tratar los datos, los cuales son demasiados actuales con una potencia superior para determinar una solución válida a la cantidad de datos registrados.

Otra propiedad de los datos adquiridos es que la Teledetección es el poder permitir dar seguimiento al desarrollo de las enormes extensiones forestales las cuales existen en la superficie de la tierra, al poder tener una perspectiva en conjunto general de las causas de los efectos que se producen debido a las grandes catástrofes como pueden ser las eternas sequías en las regiones desérticas del Sahara de África, para en dado caso poder ubicar determinados fenómenos generados por la contaminación a gran escala afectando al cielo y el ambiente en el mar.

Como el sol es una fuente de luz la cual emite radiación electromagnética en longitudes de onda los cuales van desde los 10-13 metros de los rayos gamma hasta los 105 metros de las ondas de radio largas, debido a que la radiación electromagnética solar, después de ser filtrada por la atmósfera, interacciona con la superficie terrestre por lo que es parcialmente reflejada en todas direcciones [7].

Además, se conoce que todo cuerpo el cual está a una temperatura superior a 0° Kelvin emite radiación electromagnética dentro de las frecuencias dependientes de su temperatura superficial, en el caso de la Tierra la cual es una superficie con esas características, estas frecuencias se sitúan en la banda de infrarrojo térmico, que es entre 7 y 14 micrómetros de longitud de onda aproximadamente.

Los sensores pasivos o también conocidos como ópticos miden en unidades de radiancia ($\text{W}/\text{m}^2/\text{sterad}$) la energía electromagnética solar reflejada por la Tierra, además de medir la radiación térmica emitida por la Tierra en distintas bandas espectrales o en diferentes intervalos del espectro electromagnético.

Donde la cantidad de radiación reflejada va a depender de las condiciones de iluminación presentes, más sin embargo la proporción entre la radiación

incidente y reflejada, tienen una magnitud denominada reflectancia, la cual es una propiedad específica de la superficie reflectante.

Es decir que la curva que representa la reflectancia de un material en función de la longitud de onda la cual recibe el nombre de espectro o también conocida como firma espectral del material, el cual constituye una característica identificativa del mismo material.

Por otra parte, la teledetección activa dispone de su propia fuente de radiación de microondas para poder iluminar un objetivo, en donde lo que se mide es el reflejo de esa radiación sobre tal objetivo [8].

Los sensores en la teledetección activa se pueden dividir principalmente en dos categorías las cuales se pueden separar en con imagen y "sin imagen". Debido a que los sensores activos con imagen son basados en técnicas de radar también conocidas como Radio Detection And Ranging o de radar de apertura sintética (SAR) para poder generar imágenes en dos dimensiones.

Por otro lado los sensores activos sin imagen son los que incluyen los altímetros y escatómetros, por los cuales en la mayoría de los casos son dispositivos de perfil, los cuales toman medidas en una dimensión, lo cual es lo opuesto a la representación bidimensional de los sensores de imagen.

Se llama teledetección pasiva a la técnica aplicada por los sensores que miden las variaciones de la energía procedente de los objetos sin intervenir en el campo natural, la cual se denomina teledetección activa a aquellos que generan un campo de energía artificial, utilizado para registrar y medir el efecto que en él producen los objetos, donde las características técnicas del sensor influyen en la calidad de los datos y en la posibilidad de recibir información en distintas longitudes de onda.

La preocupación generalizada de los habitantes del planeta debido a la falta de recursos naturales y falta de fuentes de energías, así como la generación de degradaciones las cuales han causado el ser humano en el medio ambiente a consecuencia de sus actividades, las cuales son en gran medida de carácter irracionales en contra de la naturaleza, lo cual ha generado en el mundo entero la necesidad de poder generar un mayor conocimiento de los espacios naturales dentro de los cuales se desarrolla la vida.

Ya que la adecuada programación de ciertas actividades las cuales se presentan en las circunstancias actuales son las que exigen que han de descansar en el poder generar una realización de un inventario más completo, en el cual se pretende actualizar el conocimiento de las riquezas naturales nacionales e internacionales, ubicadas en los rubros de las áreas agrícolas,

áreas forestales, áreas hidrológicas, áreas mineras o cualquier áreas que represente una fuente de recursos para la humanidad [9].

De tal manera que la vigilancia debe ser constante para el medio ambiente, debido a que esta actividad podrá generar una importante reducción en el impacto generado por los humanos en el medio ambiente hasta este momento, es por eso que los datos obtenidos del análisis de Teledetección se convierten en una importante fuente que genera información por lo que tiene un importante papel en el cumplimiento de los objetivos del cuidado del medio ambiente.

Es posible centrar los análisis en el caso español, considerando que las acciones más imprescindibles deben ser enfocadas en la calidad del agua, otra acción sería la detección de incendios, debido a que el agua es una gran riqueza para la Península Ibérica, básicamente es necesaria para mantener la vida y el desarrollo del ser humano, por lo que, debido a que si su calidad de la vida se deteriora, afecta a todos de tal manera que sufriríamos las consecuencias, tanto los hombres, como los animales y también las plantas.

La Teledetección aplicada en los recursos naturales está basada en un sistema el cual es capaz de adquirir datos a distancia, la cual por lo general esta ubicada sobre la biósfera, la cual está basado en las propiedades de la radiación electromagnética, junto con su interacción con los materiales que se encuentran en la superficie terrestre, ya que todos los elementos que conforman la Naturaleza son capaces de generar una respuesta espectral propia, la cual se denomina “signatura espectral”.

Por lo general la Teledetección es capaz de estudiar las variaciones espectrales, las variaciones espaciales, como las variaciones temporales de las ondas electromagnéticas, en donde se puede apreciar que las correlaciones existentes entre las diferentes variaciones, junto con las características de los variados materiales terrestres, donde su objetivo básico es el de generar la identificación de los materiales que existen en la superficie terrestre además de poder identificar ciertos fenómenos que en la naturaleza se generan por medio de su característica de espectra.

La implementación de la teledetección impacta de forma positiva en la preservación y el mejoramiento de la calidad del agua de los ríos y embalses, debido a que el uso de estas técnicas permite vigilar su situación para poder actuar y manejar cualquier vertido contaminante que se llegara a producir [10].

Además, estas técnicas de teledetección sirven de apoyo en las diversas actividades de saneamiento, ya que además con los terremotos y la actividad

volcánica se genera un gran potencial destructivo para el medio ambiente, ya que la actividad volcánica, al igual que la sísmica, son la manifestación de la liberación de energía acumulada en el interior del planeta tierra.

Por lo que los desprendimientos generados de flujos de lava, además de los fragmentos de roca incandescente, así como la presencia de gases tóxicos traen consecuencias como la destrucción de las infraestructuras existentes en los ecosistemas, generándose incendios, muerte de las plantas, asfixia de personas y animales, pudiendo ocasionar la interrupción del transporte, así como de la comunicación, sin olvidar la contaminación de las aguas.

De acuerdo con estudios existe la posibilidad de que al observarse un incremento en la comprensión de las condiciones de la corteza de la terrestre o en su interior se pueda permitir el generar un sistema de alerta temprano el cual se anticipe a tales potenciales cataclismos.

Es por ello por lo que las técnicas de teledetección impactan positivamente en la generación de un pronóstico, obteniendo la detección, el generando un monitoreo constante de semejantes desastres naturales en las diferentes áreas del planeta.

Las técnicas de teledetección impacta positivamente en la generación de alerta y monitoreos de diferentes fenómenos naturales, como la instalación de una estación de monitoreo de mareas por el Centro para Alarmas por Tsunamis en el Pacífico conocido como PTWC por sus siglas en inglés, además del Servicio Nacional de Estudios Territoriales (SNET), la cual permite obtener información sobre el nivel del mar y los cambios que se originen en las costas del Pacífico, es así como la obtención de información podría evitar daños por la generación de un Tsunami en la región de este océano la cual podría afectar a Centroamérica.

La estación de monitores está conformada por diferentes instrumentos especiales los cuales censan y recopilan la información cada dos minutos para después transmitirla cada hora vía satélite a las oficinas del PTWC, ubicadas en Hawai, en donde se recolecta y evalúan los datos proporcionados por los 28 países que realizan monitoreo de mareas, para poder diseminar boletines de alarma informativos de la ocurrencia de un sismo importante o la generación de un posible o confirmada de un tsunami.

Como ya se expresó con anterioridad, la teledetección constituye en un sistema el cual fue desarrollado para la adquisición de datos, que se encuentra a cierta distancia, ubicados sobre la biósfera, en el cual está basado en las características de la radiación electromagnética junto con su interacción con los diferentes materiales de la superficie terrestre.

Ya que todos los elementos existentes en la naturaleza cuentan con una respuesta espectral la cual es característica a la cual se la denomina "signatura espectral" la cual es capaz de estudiar las variaciones espectrales, las variaciones espaciales y las variaciones temporales de las ondas electromagnéticas, poniendo de manifiesto las relaciones existentes entre las variaciones junto con las características que tienen los diferentes materiales terrestres.

Es por ello que su objetivo básico se centra en la identificación de los diferentes materiales de la tierra y los diferentes fenómenos que en ella se generan a través de su característica espectral, desde este punto de vista, la teledetección como herramienta ofrece enormes posibilidades para la generación de avances en el conocimiento e interpretaciones de condiciones físicoambientales constituyendo una fuente de información y desempeñando un papel significativo en el campo del conocimiento geográfico [11].

Se comparte lo expresado por Goillot (1976) en cuanto a que el conjunto de datos adquiridos mediante los procedimientos de teledetección instalados en plataformas como aviones o transbordadores espaciales constituyen siempre tres tipos de información, la primera de ellas es una serie de información espacial con la cual es posible representar la organización dentro del espacio físico de los diferentes elementos que componen la imagen, la segunda es una información espectral la cual se caracteriza, pudiendo encontrar un conocimiento sobre la naturaleza de la superficie terrestre, y por último es una información temporal la cual permite la detección de los diferentes cambios operados dentro de la superficie terrestre en el transcurso del tiempo.

En este sentido la teledetección se define como el proceso de análisis de la energía reflejada por los objetos, es decir es una técnica que puede aportar una información muy valiosa para distintos campos de intervención de la Geografía, lo cual permite visualizar, entre otras aplicaciones, los distintos usos y ocupaciones del suelo, a nivel espacial y temporal, observando las transformaciones territoriales, posibles de ser plasmadas posteriormente en una carta temática elaborada en un entorno SIG.

3.3. Radar Acconeer

El radar utilizado en el proyecto está fabricado por Acconeer llamado A111. Es un radar de 60GHz basado en impulsos tecnología de radar coherente (PCR¹) totalmente integrado en un pequeño chip de 29 mm². Esto

¹Pulsed Coherent Radar

permitirá una fácil integración en cualquier dispositivo portátil impulsado por batería.

El radar A111 es un sistema de radar basado en tecnología de radar coherente pulsado también conocido como PCR el cual está estableciendo un nuevo punto de referencia para el consumo de energía y la precisión de la distancia, es totalmente integrado en un paquete pequeño de 29 mm² [12].

El sistema de radar A111 es de 60 GHz, el cual está optimizado para operar a una alta precisión, con una potencia ultra baja, por lo que se suministra como una óptima solución de un solo paquete con banda base integrada, interfaz de RF y antena en paquete (AiP), por lo que esto permitirá una fácil integración en cualquier dispositivo portátil que funcione con baterías.

El radar A111 se basa en una tecnología de sensor patentada de vanguardia con resolución de tiempo de pico de segundo, es capaz de medir distancias absolutas con precisión mm de hasta un rango de 2 m, ya que el rango de 2 m está garantizado para un tamaño de objeto, forma y propiedades dieléctricas correspondientes a un reflector esférico de esquina de 5 cm de radio, además de contar con una tasa de actualización configurable.

El radar A111 de 60 GHz no se ve afectado por ninguna fuente natural de interferencia, como lo es el ruido, el polvo, el color o la presencia de luz directa o indirecta.

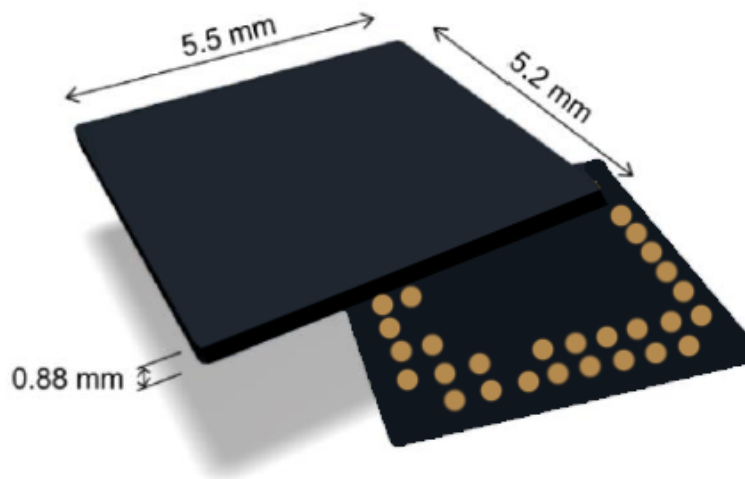


Figura 3.1: Radar A111.

Aplicaciones:

- Puede generar mediciones de distancia de alta precisión, las cuales son con precisiones de mm y alta tasa de actualización.
- Puede realizar detección de proximidad con alta precisión y posibilidad de definir múltiples zonas de proximidad.
- Puede generar detección de movimiento, detección de velocidad.
- El radar permite la detección de material.
- Se puede generar el seguimiento de objetos de alta precisión, que permite el control por gestos.
- Es posible desarrollar seguimiento de alta precisión de objetos 3D.
- Es posible que controle los signos vitales de la vida, como la respiración y el pulso.

Algunas de las características principales con las que se cuenta y que se pueden describir del radar son:

- Es un sensor completamente integrado ya que cuenta con:
 - Radar coherente pulsado (PCR) de 60 GHz
 - Banda base integrada, interfaz de RF y antena en paquete (AiP)
 - 5,5 x 5,2 x 0,88 mm fcCSP, paso de 0,5 mm
- Cuenta con un alcance y movimientos precisos de la distancia, ya que puede:
 - Medir el rango absoluto hasta 2 m o precisión absoluta en mm
 - Tiene una precisión relativa en μm
 - Cuenta con la posibilidad de reconocer movimientos y gestos para varios objetos.
 - Admite modo de barrido continuo y único
 - Tienen HPBW de 80 (plano H) y 40 grados (plano E)
- Su fácil integración es debida a que cuenta con:
 - Una solución de un chip con banda base y RF integrados

- Puede integrarse detrás de plástico o vidrio sin necesidad de una apertura física
- Cuenta con componente refluible único
- Fuente de alimentación única de 1.8 V, habilitada con Power on Reset (PoR)
- Una entrada de reloj para cristal o reloj de referencia externo, 20-80 MHz
- Una interfaz SPI para transferencia de datos, soporte de reloj SPI de hasta 50 MHz
- Soporte INTERRUPT

El radar Acconner A111 es un sensor de radar optimizado de baja potencia y alta precisión de 60 GHz con banda base integrada, una interfaz de RF y una antena en paquete (AIP).

El sensor se basa en la tecnología de radar coherente pulsado (PCR), que presenta una solución patentada de vanguardia con resolución de tiempo de picosegundos. El A111 es la elección perfecta para implementar sistemas de detección de alta precisión y resolución con bajo consumo de energía.

El radar de silicio A111 esta dividido en cuatro bloques funcionales: Power, Digital, Timing y radio mmWave.

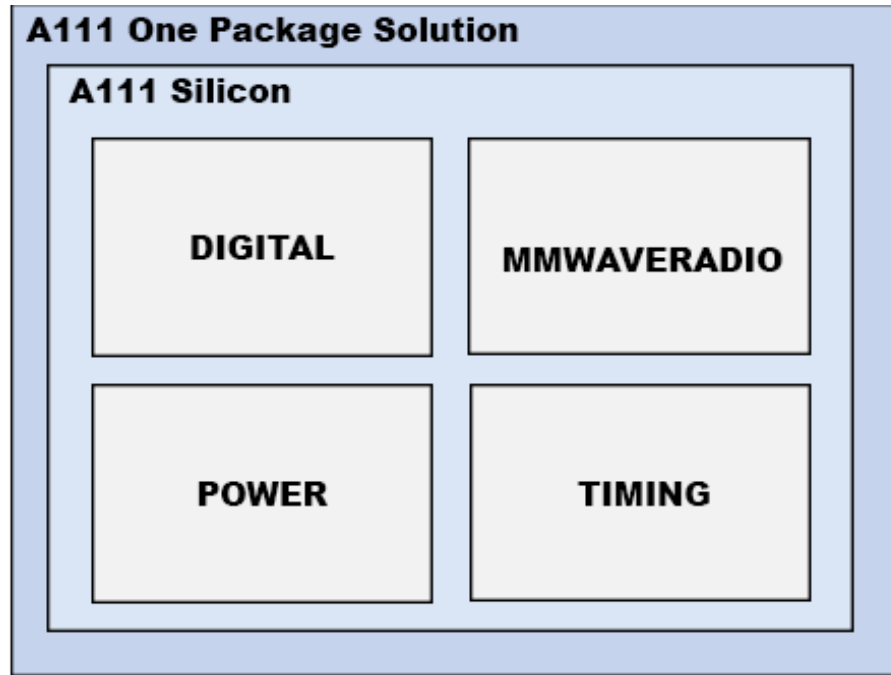


Figura 3.2: Diagrama de bloques del sensor A111.

La figura 3.2 muestra un diagrama de bloques del sensor A111. La señal se transmite desde la antena Tx y es recibida por la antena Rx, ambas integradas en la capa superior del sustrato del paquete A111. Además de la radio mmWave, el sensor consta de administración de energía y control digital, cuantificación de señales, memoria y un circuito de temporización.

El bloque funcional Power incluye LDO y un bloque Power on Reset (PoR). Cada LDO crea su propio dominio de voltaje. El bloque PoR genera una señal de reinicio en cada ciclo de encendido. El host conecta el bloque funcional de alimentación del sensor a través de una fuente de alimentación única de 1,8 V y ENABLE. El bloque funcional digital incluye control por sensor. La memoria de datos almacena los datos de barrido del radar del ADC. El host interactúa con el sensor a través de una interfaz SPI, un reloj (XIN, XOUT) y una señal de INTERRUPTIÓN.

El bloque de temporización incluye los circuitos de temporización.

El bloque funcional de radio mmWave genera y recibe pulsos de radar e incluye transmisor (TX), receptor (RX) e interfaces hacia las antenas integradas.

El software Acconeer se ha escrito en C y es portátil para cualquier sistema operativo y plataforma HW. El software Acconeer se ejecuta en Host MCU y se entrega como binarios, a excepción del software de integración que se entrega como código fuente.

El RSS (software del sistema de radar) proporciona salida en dos niveles diferentes, servicio y detector. RSS proporciona una API (interfaz de programación de aplicaciones) para la utilización de aplicaciones de varios servicios y detectores.

La salida de servicio son datos de sensor preprocesados en función de la distancia, por ejemplo. Datos de envolvente (amplitud de los datos del sensor), datos de la bandeja de potencia (datos de amplitud integrados en intervalos de rango predefinidos), datos modulados por IQ (representación en cartesiano), etc.

Los detectores se basan en datos de servicio como entrada y la salida es un resultado. Detector de distancia que presenta un resultado de distancia y amplitud basado en el servicio de envolvente, etc.

El cliente puede utilizar el detector Acconeer o desarrollar su propio procesamiento de señales basado en los datos del servicio. Acconeer proporciona varias aplicaciones de ejemplo para respaldar el desarrollo de aplicaciones propias del cliente. Además, se proporcionan pautas para el cliente para el desarrollo de aplicaciones utilizando la API RSS de Acconeer.

Acconeer proporciona varios controladores de referencia como código fuente, p. Ej. Soporte para MCU Cortex M4, Cortex M7. El software de integración implementará las funciones definidas en un archivo de definiciones proporcionado en la oferta de Acconeer Software. Esto incluye el manejo de SPI, ENABLE e INTERRUPT, así como posibles funciones del sistema operativo. Consulte la referencia HAL - Guía del usuario para obtener una guía sobre la integración de software y la implementación de HAL (<https://www.acconeer.com/products>).

La integración del espacio libre del sensor A111 debe tener en cuenta lo siguiente:

- Cualquier material por encima del sensor debe tener la permitividad y la pérdida tan bajas como sea posible, por ejemplo, plástico o vidrio con baja permitividad.
- Para concluir sobre la distancia óptima desde el sensor, se requiere una investigación de simulación / medición.

| Servicio | Tipo de dato | Ejemplo de uso |
|-----------|------------------------|--|
| Envelope | Amplitud | Distancia absoluta y presencia estática |
| IQ | Amplitud y fase | Detección de obstáculos, respiración y distancia relativa |
| Sparse | Amplitud instantánea | Velocidad, detección de presencia y detección de gestos |

Tabla 3.1: Servicios del radar A111

El sensor se puede ejecutar en uno de los siguientes servicios básicos de la tabla 3.1.

3.4. Servicio In-phase and Quadrature

El servicio In-phase and Quadrature (IQ) utiliza la coherencia de fase del radar pulsado Acconeer para producir componentes estables en fase y en cuadratura. Este servicio se puede utilizar para la detección de presencia frente al sensor, la detección de la frecuencia respiratoria, la detección de obstáculos y, en nuestro caso, para diferenciar materiales.

Los componentes en fase y en cuadratura se representan como valores complejos, lo que genera un conjunto complejo de N_D muestras representadas como $x[d]$, donde d es el índice de demora de la muestra.

Los datos obtenidos a través del servicio IQ proporcionan un método para examinar la reflectividad a diferentes distancias del sensor de radar.

Las técnicas de análisis de imágenes que permiten diferenciar partes de las misma se han venido desarrollando debido al gran número aplicaciones en donde se emplean, donde la detección y el análisis temporal de cambios sobre la superficie terrestre donde también es posible el mapeado de inundaciones, mediciones de campos, determinación de extensión de hielos continentales y antárticos, flujos de hielo de glaciares, etc., debido a que es un área de gran interés.

En este sentido, se ha venido trabajando en el análisis de las imágenes satelitales obtenidas por Radares de Apertura Sintética conocidas como SAR, para poder generar el análisis de las imágenes se usan técnicas de contraste de características estadísticas mediante las cuales se producen

mapas diferenciales que permiten identificar las regiones en las que se produjeron variaciones [13].

Es posible el aplicar algoritmos en los que mediante la definición de umbrales adecuados se diferencian las regiones, estableciéndose la presencia de bordes, por ejemplo, usando la diferencia entre los valores entre pixeles, o en el cociente entre los valores de los pixeles.

Y que dado los sistemas de adquisición de SAR basan su capacidad de formar imágenes en que se mantiene la coherencia entre las señales emitida y recibida a lo largo de toda una adquisición, la coherencia de estas señales permite, mediante el proceso de enfocado, obtener imágenes con una resolución mucho mejor que la esperada por el IFOV (Instantaneous Field of View) de la antena.

El problema que se presenta en este método es la aparición de un ruido multiplicativo conocido como speckle, en donde este ruido se debe a la interferencia que se produce al sumar las señales coherentes provenientes de los elementos dispersores presentes en la región asociada al pixel considerado.

Por lo que se hace necesario el uso de herramientas matemáticas más sofisticadas para la detección de bordes, donde las imágenes SAR son ofrecidas en distintos formatos y niveles de procesamiento, siendo la más básica de las imágenes enfocadas la denominada SLC (Single Look Complex).

Una adquisición SAR en polarización simple, vertical u horizontal, tiene dos bandas de datos que corresponden a los canales de adquisición I (in phase) y Q (in quadrature), los cuales pueden ser interpretados como la parte real e imaginaria, respectivamente, de una señal compleja.

Donde el speckle ha sido descripto teóricamente mediante modelos de distribución en regiones homogéneas, donde en estos modelos establecen que para imágenes SLC de polarización simple correspondientes a zonas homogéneas (por ejemplo áreas rurales), los datos de ambos canales, I y Q, tienen una distribución Gaussiana para las amplitudes.

Mientras que la fase tiene una distribución prácticamente Uniforme y la intensidad de una distribución exponencial (o equivalentemente Gamma de orden 1), mientras que el módulo de la amplitud (que es proporcional a la raíz cuadrada de la intensidad) responde a una distribución de Rayleigh .

Considerando esto, se han usado diferentes técnicas que tienen en cuenta la correspondiente distribución de la potencia que se recibe de cada pixel y se han calculado las distribuciones para los gradientes de intensidad asociados a los pixeles, con esto se ha conseguido una mejor identificación de los bordes

Para dar servicio a las diferentes misiones interplanetarias de la agencia espacial europea, conocida como ESA, las estaciones disponen de un receptor de frecuencia intermedia (IFMS por sus siglas en inglés), que es utilizado para múltiples propósitos como ser la transmisión de telecomandos, la recepción de telemetría, la decodificación de datos y las mediciones Doppler de las señales [14].

La estación permite la recepción de datos en un modo denominado de bucle abierto mejorado (EOLP), por lo que en este modo, el IFMS puede procesar 2 bloques de banda ancha del espectro de frecuencia intermedia (70 MHz), en estos bloques pueden corresponder a diferentes polarizaciones centradas en una misma frecuencia o la misma polarización centrados en diferentes frecuencias.

Ya que pueden capturarse hasta 8 subcanales centrados a diferentes frecuencias, dependiendo de la configuración de la estación, en el ancho de banda de los canales puede configurarse en un rango de 1 kHz a 2 MHz, ya que también puede configurarse la cantidad de bits de cada muestra lo que se conoce como cuantización de 1 a 16 bits.

Para las observaciones astronómicas se utilizaron los tres IFMS disponibles que se utilizaron de manera tal de abarcar el mayor ancho de banda de frecuencia posible, como la mayoría de los equipos de telecomunicaciones, la antena divide el ancho de banda espectral total en canales de frecuencias (12 subbandas en total, en este caso) y realiza una conversión de la frecuencia de banda X (8 GHz) o Ka (32 GHz) a frecuencia intermedias.

La información de la señal adquirida en la observación se almacena en las unidades de almacenamiento externo (ESU) en archivos que contienen paquetes UDP/IP, cada uno con una cabecera y un bloque de datos crudos compuestos por muestras de valores complejos conocidos como (In-phase and Quadrature, IQ).

Un radar Doppler tiene etapas de detección síncrona, es decir, que están en sincronía con el oscilador de transmisión, lo que define un receptor coherente, debido a que las señales Doppler contienen no sólo información de amplitud sino también de fase, es crucial conocer estas componentes para una correcta demodulación.

Una forma de obtenerlas es indirectamente, conociendo la amplitud de la señal en fase (I) y la amplitud de la señal en cuadratura (Q), donde estas señales se obtienen al mezclar la señal recibida con las señales obtenidas de un oscilador sincronizado que genere una en fase y la otra desfasada 90° [15].

Una forma de obtener estas componentes es utilizando receptores enteramente digitales, aunque el receptor analógico tuvo muchos años de vigencia, las técnicas actuales permiten realizar receptores enteramente digitales en un sólo chip.

Es importante notar que la detección de la información se puede realizar en banda base o en una frecuencia intermedia, debido a que la señal de la antena y la señal del oscilador local tendrían la misma frecuencia en el caso del receptor homodino.

Un receptor digital consiste en una fuente analógica, en este caso la señal recibida por la antena, la cual es previamente es filtrada y amplificada por un amplificador de bajo ruido (LNA), para a continuación la señal es digitalizada por un conversor análogo-digital (ADC) y luego es multiplicada digitalmente por una señal de referencia para obtener la frecuencia intermedia o la señal en banda base.

Además de la señal de entrada RF se obtienen dos señales, la primera es producto de la multiplicación por la referencia, y la otra es producto de la multiplicación por la referencia desfasada 90° , donde el objetivo es obtener información no sólo de magnitud sino también de fase respecto a la referencia.

Cabe señalar que el receptor digital concentra su selectividad en el filtro digital, donde este filtro, para el caso del radar Doppler, suele ser un filtro adaptado (matched filter) con el fin de obtener la mejor relación de señal a ruido.

3.5. Aprendizaje automático

El aprendizaje automático o machine learning es un tipo de inteligencia artificial (AI), consiste en programar una computadora para que mejore en la realización de una tarea a partir de datos de ejemplo o de la experiencia.

La inteligencia artificial es un concepto de creación de máquinas inteligentes que estimula el comportamiento humano, mientras que el aprendizaje automático es un subconjunto de la inteligencia artificial que permite que la máquina aprenda de los datos sin ser programada.

La diferencia entre el software informático normal y el aprendizaje automático es que un desarrollador humano no ha dado códigos que le indiquen al sistema cómo reaccionar ante la situación, sino que está siendo entrenado por una gran cantidad de datos.

Con la cantidad de la información la cual es generada diario en Internet, ya sea por medio de las redes sociales, por medio de diferentes transacciones comerciales o datos generados por distintos dispositivos, entre otros, existen diferentes procesos los cuales aprovechan toda esa información, en lugar de almacenarla como data la cual solo está ocupando demasiado espacio en los servidores, estos datos se continúan empleando para poder generar un análisis el cual establezca comportamientos para poder identificar las tendencias futuras [16].

En ciertas ocasiones se reúne demasiada información, por lo que es posible determinar con anticipación, de forma segura el comportamiento futuro de un grupo de personas los cuales interactúan con equipos electrónicos, ya que conoceremos sobre la técnica del Machine Learning, elemento fundamental de la Ciencia de Datos, los métodos que utiliza para realizar las predicciones de datos y su presentación.

Los diferentes dispositivos que cuentan con la conocida inteligencia artificial, son capaces de ejecutar distintos procedimientos análogos con el comportamiento humano, como es el caso de una devolución a una respuesta por cada petición de entrada parecido o similar a un tipo de reflejos que se encuentran en los seres vivos, estableciendo un estado específico entre los diferentes estados posibles según una determinada acción, con la solución de problemas empleando una lógica formal.

Cuando se les es otorgada a un determinado tipo de dispositivos la habilidad de poder aprender, se genera la posibilidad de que puedan discernir, por lo tanto se convierte en entidades que razonan con capacidades semejantes a las de un superhombre, dado que su velocidad de procesamiento es prácticamente imposibles para un humano promedio, las maquinas no tienen la necesidad de descansar para poder desarrollar sus funciones correctamente, son algunas de entre muchas grandes ventajas que ubican a las maquinas sobre los seres vivos en este contexto.

En este sentido es posible encontrar tres grupos de algoritmos en Machine Learning, los cuales son los algoritmos supervisados, donde son los algoritmos que emplean un conjunto de datos los cuales son de entrenamiento etiquetados o también conocidos como preclasificados, los datos son procesados para poder realizar predicciones sobre ellos, existiendo la posibilidad de corrección cuando son incorrectas [17].

Otro grupo es el llamado proceso de entrenamiento, el cual continúa hasta donde el modelo alcanza su nivel deseado en cuanto a su precisión, además de considerar del grupo de algoritmos semi-supervisados, estos algoritmos combinan tanto datos etiquetados como datos no etiquetados para poder

generar una función determinada o clasificada, donde estos tipos de modelos son los que tienen que aprender las estructuras para poder organizar los datos, es así como pueden realizar predicciones.

Otro grupo es el de los algoritmos no supervisados los cuales son el conjunto de datos los cuales no se encuentra etiquetado, por lo que no se cuenta con un resultado conocido, ya que por ello es necesario deducir las estructuras que se encuentran presentes en los datos de entrada, lo cual puede ser conseguido a través de un proceso matemático para poder bajar la redundancia sistemáticamente, tratando de organizar los datos por similitud.

Dentro de esta clasificación se pueden encontrar diferentes números de algoritmos específicos con las diferentes características para poder generar el tratamiento de los datos, entre los algoritmos más relevantes podemos encontrar el de Deep Learning, el cual consiste en el empleo de algoritmos para poder hacer representaciones abstractas sobre la información, para facilitar el aprendizaje automático.

Otro algoritmo es Active Learning, el cual es un caso de aprendizaje semi-supervisado en el cual el algoritmo de aprendizaje puede interactuar con un usuario, además de otra fuente de información para poder obtener los resultados deseados, otro algoritmo es el Support Vector Machines, es un algoritmo el cual busca generar la maximización de la distancia entre la recta, también conocida como plano y las diferentes muestras las cuales se encuentran de un lado u otro, ya que en el caso que las muestras obtenidas no sean linealmente separables se emplea una transformación llamada kernel.

El aprendizaje de máquina es un área que estudia cómo construir programas de computadoras que mejoren su desempeño en alguna tarea gracias a la experiencia, debido a que esta se basa en las ideas de diversas disciplinas, como inteligencia artificial, estadística y probabilidad, teoría de la información, psicología y neurobiología, teoría de control y complejidad computacional.

El Machine Learning es la capacidad que posee un software o una máquina para aprender a través de la aplicación de varios algoritmos de su programación, es decir identificar modelos complejos en múltiples datos, además de que el Aprendizaje Automático es una tecnología avanzada que permite que muchas operaciones se puedan realizar simultáneamente reduciendo la necesidad de intervención humana [18].

Este sistema proporciona una gran ventaja a la hora de examinar de manera más efectiva una mayor cantidad de datos, como podemos observar las tres fuentes de investigación nos indican que el aprendizaje automático

(Machine Learning) es de mucha importancia ya que en la actualidad debido a los avances tecnológicos que existen se han creado nuevas experiencias tecnológicas en diferentes áreas tales como: empresas, educación, sociedad, entre otros.

El aprendizaje automático es el entrenar una máquina para que aprenda a través de los algoritmos esta tecnología ayuda a la automatización de muchos procesos, lo cual ayuda a que los procesos tengan un bajo índice de error.

El machine learning se puede visualizar en diferentes áreas como la inteligencia artificial, procesos estadísticos, entre otros, ya que el aprendizaje automático tiene muchas ventajas ya que reduce el tiempo de demora de los procesos y menos personal para realizar las actividades presenciales.

El aprendizaje computacional (machine learning en inglés) estudia la construcción de sistemas capaces de aprender a partir de datos, lo que esto incluye una gran variedad de sistemas, desde sistemas de visión por ordenador hasta sistemas para detectar correo no deseado (spam).

En todos los casos, un sistema que aprende debe ser capaz de generalizar, es decir, de encontrar patrones y regularidades en los datos que le permitan desempeñarse bien en datos que no ha observado previamente, ya que existen dos tipos principales de modelos de aprendizaje computacional, modelos de aprendizaje supervisado y modelos de aprendizaje no supervisado.

En el aprendizaje supervisado se busca inducir modelos capaces de predecir el valor de ciertas variables dependientes a partir de variables independientes, un ejemplo de problema de aprendizaje supervisado es el problema de clasificación, en el cual la variable dependiente corresponde a un atributo que indica a qué clase pertenece, por ejemplo, enfermo o control en el caso de un problema de diagnóstico médico, el cual pertenece una muestra particular.

En los modelos de aprendizaje no supervisado no hay una distinción entre variables dependientes y no dependientes, en este caso se pretende encontrar la estructura subyacente que explique la estructura de los datos [19].

Y que el ejemplo más representativo de aprendizaje no supervisado es el análisis de conglomerados también conocido como clustering en inglés, en el cual el objetivo es encontrar grupos de datos que compartan características similares.

La minería de datos se centra en desarrollar y aplicar algoritmos que, bajo limitaciones aceptables de eficiencia computacional, permitan la obtención de patrones sobre los datos, los cuales buscan esta obtención de patrones que pueden abordarse mediante dos objetivos básicos.

La predicción es el sistema el cual busca patrones para predecir un comportamiento futuro, de los modelos desarrollados dentro de este objetivo están basados en aprendizaje supervisado, de entre las tareas predictivas se pueden encontrar, por ejemplo, la clasificación, la regresión o el análisis de series temporales.

La descripción, es donde el sistema busca patrones para presentarlos a un experto en una forma comprensible para él, y que describen y aportan información de interés sobre el problema y el modelo que subyace bajo los datos. Los modelos desarrollados en este enfoque se basan en aprendizaje no supervisado, entre las tareas descriptivas se pueden encontrar, por ejemplo, el agrupamiento, la sumarización o la asociación.

A pesar de existir una clara distinción entre las técnicas dependiendo de los objetivos que pretenden abordar, en la actualidad existen un conjunto de técnicas que se encuentran a medio camino entre las técnicas predictivas y las descriptivas, agrupadas bajo el nombre de descubrimiento de reglas descriptivas basadas en aprendizaje supervisado, que intentan obtener reglas o conjuntos de ítems de una categoría o clase prefijada para describir información significativa y relevante del conjunto de datos.

Su principal objetivo no es clasificar nuevas instancias, sino comprender o encontrar fenómenos subyacentes; es decir, encontrar información desconocida u oculta difícil de descubrir por los expertos, sobre el valor de una clase prefijada o variable de intereses.

Dentro de este grupo de técnicas se encuentran la: minería de conjuntos de contraste, el descubrimiento de subgrupos y los patrones emergentes, donde la minería de patrones emergentes busca conocimiento relacionado con los distintos valores de la variable objetivo, donde el número de instancias cubiertas por un patrón sea muy elevado para un valor de la variable objetivo y muy bajo (o cero) para el resto; es decir, que el mismo patrón tenga un soporte muy alto para una clase y muy bajo para las demás [20].

Tipos de aprendizaje automático

Aprendizaje supervisado

El algoritmo de aprendizaje conocido como supervisado emplea un programa capaz de recibir datos de entrada etiquetados y los datos de salida esperados. Obtiene los datos de los datos de entrenamiento que contienen conjuntos de ejemplos.

Generan dos tipos de resultados:

- Clasificación: Notifican la clase de los datos que se presentan.
- Regresión: esperan que el producto produzca un valor numérico.

El aprendizaje supervisado, se da cuando se entrena un algoritmo de Machine Learning aportándole las preguntas, las cuales son características, uno con una serie de respuestas conocidas también como etiquetas, por lo que así en el futuro el algoritmo sea capaz de hacer una predicción bajo el conocimiento de las características, por lo general en este tipo de aprendizaje hay dos algoritmos, llamados entrenamientos, los cuales son el algoritmo de clasificación y el algoritmo de regresión.

Los algoritmos de clasificación son los que esperamos que el por su naturaleza el algoritmo nos indique a qué grupo pertenece determinado elemento en estudio, por lo general el algoritmo tiene la capacidad de encontrar patrones en los datos que le proporcionamos pudiendo clasificarlos en grupos, ya que luego se comparan con nuevos datos obtenidos para después poder ubicarlos en uno de los grupos, siendo de esta manera como puede predecir de que se trata de objeto de estudio.

Aprendizaje sin supervisión

El aprendizaje sin supervisión es un tipo de algoritmo el cual consta de datos de entrada sin etiquetar, el cual esta caracterizado por su intervención humana menor, en este algoritmo por lo general se emplea principalmente en análisis exploratorios, debido a que puede identificar automáticamente la estructura en los datos.

En el caso de los algoritmos de aprendizaje no supervisado, no es necesaria de la intervención humana para poder desarrollar un conjunto de datos los cuales son previamente clasificados para poder representar el algoritmo de aprendizaje, dado que el objetivo de la enseñanza no supervisada es el poder encontrar modelos interesantes teniendo en cuenta la distribución con la

compilación de los datos que se presentan, algunos ejemplos que se pueden mencionar sobre técnicas de aprendizaje no supervisado son las técnicas que se aplican de agrupación [?].

Aprendizaje semisupervisado

El algoritmo de aprendizaje semisupervisado emplea el método del aprendizaje que es automático el cual evita la gran cantidad de documentos etiquetados mejorando este tipo de limitación ya que aprende a clasificar a partir de un número casos de ejemplos etiquetados junto con otros no etiquetados, dado que los ejemplos etiquetados se usan para poder aprender modelos los cuales caractericen cada clase o categoría, en donde los ejemplos sin etiqueta se usan para refinar los límites entre las clases [?].

El llamado Self-training la cual es una de las formas más simples de clasificación semi supervisada en donde primero se construye un clasificador usando los datos etiquetados, para después emplear el clasificador el cual categoriza a los datos no etiquetados, ya que sólo los nuevos ejemplos etiquetados con una confianza que supere cierto umbral se anexan al conjunto etiquetado, es cuando el proceso de aprendizaje se repite.

Dado que este tipo de algoritmos semi supervisados empleen pocas instancias etiquetadas para poder entrenar hace que sean atractivos de emplear en los casos donde el etiquetado sea muy costoso, donde se requiera tiempo o donde se requiera intervención humana, en diferentes proyectos donde emplearon los algoritmos semi supervisados, se evalúan con respecto a una serie de algoritmos semi supervisados los cuales son aplicados a textos de micro-blogs por lo general en idioma inglés, dado que este enfoque es muy importante para la clasificación textual.

Por obvias razones los algoritmos de aprendizaje semi supervisado es un término el cual se encuentra entre el aprendizaje supervisado y el aprendizaje sin supervisión, dado que en los algoritmos de aprendizaje semi supervisado, es posible dividir los datos en dos partes las cuales están integradas por un grupo de datos clasificados y también está integrada por un grupo de datos no clasificados, por lo consiguiente se llama aprendizaje semi supervisado estándar.

Para diferentes investigadores el aprendizaje semi supervisado es más útil cuando hay más datos no clasificados comparado con los datos clasificados, dado que especialmente cuando es necesario mucho esfuerzo para poder obtener datos clasificados, en donde tiende a tardar mucho tiempo aumentando

el costo, debido a que obtener datos no clasificados generalmente es más barato.

Algunos ejemplos donde se aplican las técnicas de aprendizaje semi supervisado es en las técnicas de máquinas de vectores de soporte transductivo, otro ejemplo es la maximización de las expectativas, además de poder mencionar algunas aplicaciones donde se aplican las técnicas de aprendizaje semi supervisado mencionadas las cuales podrían ser la clasificación de páginas web, el reconocimiento de voz y la secuencia de la proteína.

Aprendizaje reforzado

El algoritmo de aprendizaje reforzado interactúa con el entorno, dado que es posible emplearlo para tomar una secuencia de decisiones, por lo general se puede plantear como un método de seguimiento en el cual es posible encontrar un error lo suficientemente apto para encontrar el mejor resultado basado en la experiencia.

El aprendizaje reforzado es un campo dentro del aprendizaje automático, dado que los problemas los cuales se tratan en esta área requieren de la toma de un conjunto de decisiones secuenciales para poder conseguir un objetivo determinado, dado que, en estas situaciones, una entidad denominada “agente” es la que interactúa con el entorno debido a que aprende de forma autónoma las acciones que debe llevar a cabo con el fin de maximizar una señal numérica escalar conocida como recompensa [?].

Por lo general, en un principio el agente no conoce la dinámica del entorno en el que se encuentra, tampoco conoce los detalles explícitos de la tarea que debe realizar, debido a que la única vía de la que dispone el agente para aprender la dinámica, con los detalles de tal forma que obtenga la máxima recompensa es mediante la prueba y el error, dicho de otra manera, la experimentación es la única forma de la que se dispone para poder desarrollar el entrenamiento.

Dado que la finalidad de un agente al resolver un problema de aprendizaje reforzado es la de aprender el comportamiento el cual le permitirá alcanzar una recompensa óptima, dado que el comportamiento de un agente queda determinado mediante una política de actuación, considerando que una característica peculiar de los problemas que se presentan en el aprendizaje reforzado es que la recompensa no tiene por qué ser inmediata.

Dado que la recompensa para el agente puede llegar después de haber sido ejecutada en conjunto con una serie de acciones las cuales no tienen por qué ser óptimas individualmente, dado que los agentes no deben actuar

de forma voraz, dado que la mejor opción localmente no siempre lleva a una buena recompensa a largo plazo.

Dado que las dos cualidades implementadas como la experimentación y la recompensa con retraso forman parte esencial de los problemas de aprendizaje reforzado, son estas las cualidades distintivas las cuales no pertenecen a ningún otro campo del aprendizaje automático.

En el aprendizaje reforzado, el algoritmo puede aprender observando el mundo que lo rodea, dado que la información de entrada es la retroalimentación que recibe del mundo exterior en respuesta a sus acciones, por lo que el sistema aprende sobre la base de pruebas y errores, en donde en lugar de que un instructor que le indique al agente qué hacer, el agente inteligente debe aprender cómo es que el medio ambiente se comporta a través de la recompensa conocida como refuerzos o también conocido como castigos, el cual es el resultado del éxito o fracaso.

Se puede decir que el objetivo de este aprendizaje es el poder aprender la función de valor la cual ayuda al agente inteligente para poder maximizar la señal de recompensa buscando optimizar sus políticas para comprender el comportamiento del entorno para poder tomar las decisiones correctas logrando así cumplir con sus objetivos formales, donde los principales algoritmos de aprendizaje por refuerzo son desarrollados como parte de los métodos de resolución de problemas de decisión finitos de Markov, los cuales incluyen ecuaciones Bellman además de las funciones de valor, dado que los tres métodos principales son el métodos de Monte Carlo, la Programación Dinámica y el aprendizaje de Diferencias Temporales.

Podemos decir que entre las implementaciones desarrolladas podemos encontrar AlphaGo, el cual es un programa de Inteligencia Artificial desarrollado por Google DeepMind para poder jugar el juego de mesa Go, podría resultar interesante mencionar que en marzo de 2016, AlphaGo ganó el partido ante el jugador profesional Lee Se-Dol, el cual tiene la categoría novena dan con 18 títulos mundiales, además de que entre los algoritmos utilizados podemos encontrar el árbol de búsqueda de Monte Carlo, el cual también utiliza el aprendizaje profundo con redes neuronales.

Transducción

El sistema de transducción no crea un modelo para poder replicar los resultados como el método inductivo-deductivo, en el sistema de transducción emplea los datos proporcionados en la entrada para obtener los resultados para poder generar nuevos datos, dado que si existen varios grupos bien

definidos, pero de los cuales solo conocemos la clase de unos pocos, ya que el número de ejemplos que tenemos es muy reducido como para poder aplicar aprendizaje supervisado, más sin embargo podemos emplear un algoritmo de K-Vecinos Cercanos también conocido como KNN para poder obtener la clase resultante de los datos desconocidos sin necesidad de crear un modelo de inferencia [?].

La transducción es muy semejante al aprendizaje supervisado, más sin embargo no constituye explícitamente una función, ya que los datos no tienen una etiqueta, por lo cual los datos son sin clasificar, la transducción esta diseñada para predecir las categorías de los ejemplos futuros los cuales están basados en sus categorías, además de estar basados en los ejemplos de entrada y estar basados en los nuevos ejemplos en el sistema.

Aprendizaje multitarea

El aprendizaje multitarea, también conocido como Multitask Learning (MTL), esta basado en la resolución de varias tareas al mismo tiempo, el cual está aprendiendo tanto de las características comunes como de las diferencias entre ellas, por lo cual el modelo aprende en paralelo todas las tareas, al mismo tiempo comparte la información aprendida con cada una de las tareas para poder resolver las otras tareas, por lo general esta técnica es especialmente útil cuando las tareas están relacionadas, además de estar poco muestreadas, aunque también se ha demostrado que puede ser muy poderoso para tareas no relacionadas [?].

El aprendizaje multitarea es la aplicación de los métodos de aprendizaje los cuales emplean el conocimiento previamente aprendido por el sistema, los cuales se emplean cuando se enfrentan a problemas similares a los ya vistos, lo que implica generar una solución simultánea de diferentes tareas, dado que el aprendizaje de la tarea se mejora para poder complementarla con el aprendizaje común de otras tareas relacionadas con la primera tarea.

Unas de las ventajas más importantes al emplear aprendizaje multitarea, comparado con las alternativas más comunes de reglas comerciales codificadas, análisis manual y modelos estadísticos simples son las que comprenden la automatización, la precisión, la personalización, la rapidez, y la escalabilidad.

3.6. Algoritmos machine learning

Machine Learning es una disciplina científica la cual se enfoca en crear sistemas que aprenden de datos los cuales son capaces de realizar predicciones e identificar patrones complejos en ellos, se puede establecer que por definición estos sistemas tienden a mejorar de forma autónoma, el científico informático Tom Mitchell establece una definición moderna de lo que es Machine Learning, la cual se describe como, “Se dice que un programa de ordenador aprende de la experiencia E con respecto a alguna clase de tareas T y la medida de rendimiento P , si su desempeño en tareas en T , medido por P , mejora con la experiencia E .”, dada esa definición es posible identificar dos tipos de aprendizajes, los cuales son el aprendizaje supervisado y aprendizaje no supervisado [?].

La principal diferencia entre ambos está en el conjunto de datos que usaremos para entrenar, si usamos datos etiquetados o con un valor real estaremos hablando de aprendizaje supervisado mientras que, si el conjunto de datos no está etiquetado, nuestro algoritmo deberá buscar un patrón en estos, por lo que estaremos ante un problema de aprendizaje no supervisado.

Los lenguajes de programación más usados son Python, R, C++, Java y Scala. R se usa debido a que es un lenguaje muy orientado al análisis estadístico y cuenta con muchas librerías desarrolladas las cuales pueden analizar un conjunto de datos, para obtener modelos, dado que, a su vez, Python destaca por su sintaxis intuitiva y una gran variedad de librerías para el desarrollo de aplicaciones de Machine Learning, las más relevantes son:

- [•] Pandas: Nos ofrece una forma versátil de recoger estos datos en streaming para poder procesarlos posteriormente en dataframes.
- SciPy: Herramientas de matemáticas, ciencia e ingeniería.
- Statsmodel: Nos provee herramientas dedicadas para el análisis de series temporales.
- Scikit-learn: Para algoritmos de Machine Learning.
- Numpy: Para vectores y matrices.
- TensorFlow: Para integración de redes neuronales.
- Keras: Al igual que TensorFlow, es una librería destinada al desarrollo de redes neuronales.

Los algoritmos que se han propuesto para poder ser utilizado en el planteamiento del desarrollo del proyecto son Random Forest, Regresión Logística, y KNN, los cuales se describirán a continuación.

Random Forest

Para poder establecer un algoritmo basado en Random Forest, es necesario describir en qué consiste un árbol de decisión, ya que en él podemos encontrar la combinación de los diferentes casos.

El árbol de decisión es una herramienta de aprendizaje supervisado más implementadas, este modelo consta de una serie de nodos, en los cuales se toman decisiones lógicas de forma secuencial, este se presenta en diferentes ejemplos, si quisiéramos conocer si a alguien le deberíamos conceder un crédito, podríamos ver si en primer lugar obtiene más o menos de X ingresos [?].

En dado caso que fuera afirmativo podríamos querer saber si lleva más de Y años en su trabajo, o si tiene algún aval, dado que un árbol de decisión no es más que una acumulación de este tipo de decisiones, donde empezamos en el nodo conocido como raíz, pasamos por los nodos intermedios y nuestro resultado es proporcionado por un nodo terminal, que serían las hojas del árbol.

Al final, los árboles de decisión lo que hacen es atribuir un valor constante al output dentro de regiones que suelen ser rectángulos definidos en el espacio de los inputs, un ejemplo, se puede evaluar cuando se define un valor crítico (threshold) la cual es una variable, en donde se elige una frontera para tomar una decisión u otra, donde se pregunta cuál es el mejor predictor para valores superiores al valor crítico y otro para los inferiores, y esto se hace luego con otra variable, lo que se hace es particionar el dominio de los inputs dividido en cuatro pedazos y elegir el predictor constante dentro de cada uno de los pedazos. La ventaja principal de estos métodos en la sencillez de interpretación, ya que lo podemos encontrar un caso con dos variables, donde el dominio de la variable X1 se ha dividido en tres regiones, mientras que el de X2 se descompone en función del valor de la primera variable. Random Forest conocido en castellano como "Bosques Aleatorios, a cual es un conjunto de árboles los cuales predicen, por lo que la información generada en un árbol es respuesta directa de los diferentes valores obtenidos en un vector generado como muestra de manera separada con las mismas características para los demás árboles del conjunto.

Los árboles se crean siguiendo el algoritmo:

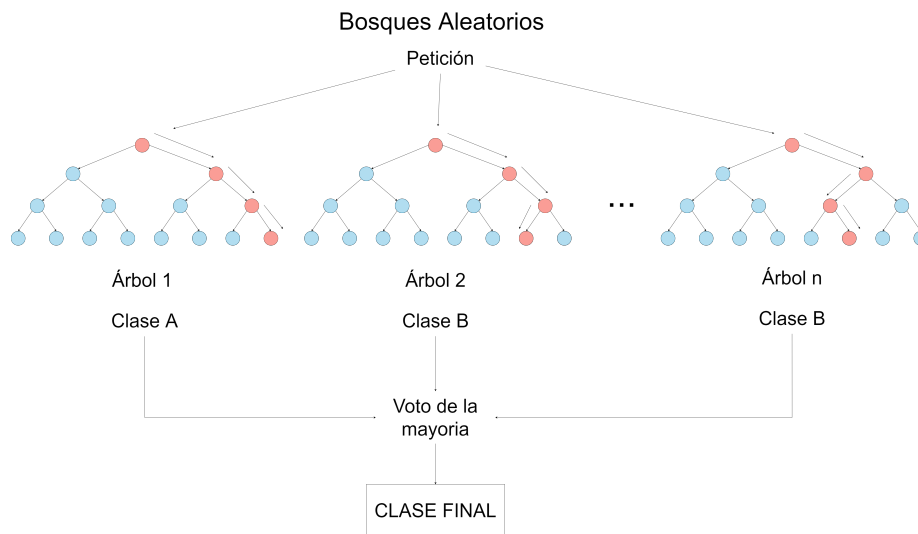


Figura 3.3: Random Forest

- Sea N el número de casos de prueba, M es el número de variables en el clasificador.
- Sea m el número de variables de entrada, m menor que M
- Se elige un conjunto de datos para el entrenamiento del árbol y el resto de los casos se utilizará para estimar el error.
- Para cada nodo del árbol, es posible elegir aleatoriamente m variables en las cuales basar la decisión, para poder calcular la mejor partición del conjunto de entrenamiento a partir de las m variables.

La idea es que a medida que se vaya separando la muestra de entrenamiento a través de los diferentes nodos, se vayan formando grupos cada vez más homogéneos, hasta que las decisiones no generen grupos más uniformes, esto debido a que el método consiste en fragmentar el espacio formado por los datos obtenidos asignando un valor como resultado para todos los subespacios, nada impide que este método sea capaz de capturar relaciones no lineales.

Por las cualidades del método, además de considerar la naturaleza del proyecto presentado en este trabajo el algoritmo Random Forest ha sido el seleccionado para ser implementado para poder realizar el proyecto.

Regresión Logística

El método de regresión logística es uno de los métodos estadísticos más utilizados para resolver problemas de clasificación binaria lo cual se refiere a la clasificación de dos clases, donde el resultado solo puede ser de naturaleza dicotómica, o sea, solo puede tomar dos valores posibles.

Como técnica de Machine Learning, la regresión logística se le puede definir como: *“Es una regresión enmarcada en el grupo de los modelos lineales generalizados, las cuales utilizan una función logit y son útiles para modelar probabilidades referentes a un evento en función de otras variables, se basa en la probabilidad de ocurrencia de los posibles valores de la variable dependiente y utiliza para ello una regresión logística que acota los resultados en el intervalo 0 y 1, adicionalmente, está respaldada por su capacidad de pronóstico.”* [?].

Es posible aplicar la regresión logística siempre y cuando las variables sean categóricas, siendo aplicado el modelo para poder calcular los diferentes valores de la variable que se estudia, por lo que se emplea el cálculo de la probabilidad en donde la variable en la que se enfoca el estudio tiene el valor de los eventos que se definieron anteriormente.

La variable independiente que se analiza puede tomar diferentes valores por lo que no tiene restricciones de rangos de valores, se recurre a métodos de estimación aplicados a modelos de regresión para elaborar un modelo de regresión logística, siendo que la regresión logística es un modelo que se obtiene empleando diferentes modelos matemáticos semejantes, después de realizar transformaciones empleando los datos.

Este algoritmo emplea para solucionar varios problemas de clasificación, detección de spam, predicción de la diabetes, si un cliente determinado comprará o no un producto en particular.

KNN

KNN también conocido como k-nearest neighbors, siendo un algoritmo de tipo simple en el cual se puede almacenar los casos disponibles, clasificando los casos nuevos, se basa en un control de parecido, es posible analizar funciones de longitudes, donde KNN se emplea para obtener el reconocimiento de patrones, también es empleado para obtener la estimación estadística, ya que es una técnica no paramétrica, no hace suposiciones con los datos subyacentes.

K-nearest neighbors traducido en castellano como "k vecinos más próximos" su funcionalidad es poder buscar un número anteriormente definido de un conjunto de datos de entrenamiento la cual estará cercana al nuevo punto, pudiendo predecir la etiqueta a partir de los datos.

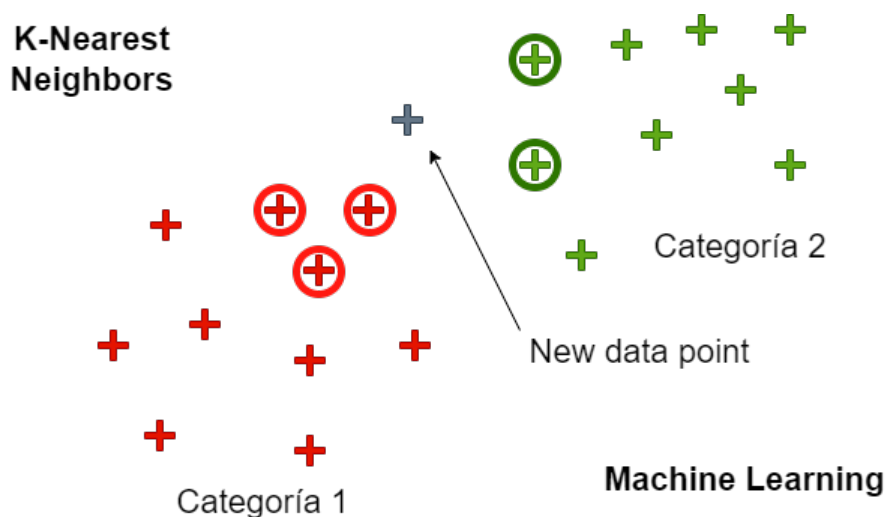


Figura 3.4: K-nearest neighbors

La cantidad de datos son una constante la cual es definida por los usuarios conocida como aprendizaje del vecino más cercano k , es posible que varíe de acuerdo con la cantidad local de puntos, donde la distancia es cualquier medida, por lo que la distancia euclidiana estándar es la más empleada. Dado que el dato más cerca de K es un algoritmo de aprendizaje automático basado en técnicas de aprendizaje supervisado, se asume el parecido entre el nuevo caso con los datos, ya que los casos disponibles para poder establecer el nuevo caso en su categoría deberá ser más similar a las categorías existentes [?].

El algoritmo KNN es capaz de almacenar los datos disponibles, clasificando un punto en función de la similitud, por lo que cuando aparecen datos nuevos, es posible clasificar fácil en una categoría de conjunto de datos empleando el algoritmo KNN, siendo empleados para la regresión, la clasificación y problemas de clasificación.

Conocido como algoritmo de aprendizaje perezoso ya que guarda el conjunto de datos, para clasificarlos en una acción en el conjunto de los datos, en donde los algoritmos KNN cuentan con dos parámetros, la K que representa el número de vecinos y la k que es un parámetro de suavizado,

donde mientras más grande sea k más pequeño es el ruido, donde d es a cantidad de retardos para cada elemento.

3.7. Características de los materiales

Se ha realizado una extracción de características de los materiales utilizados para el entrenamiento del modelo de clasificación, en donde las características y además atributos obtenidos los cuales son una serie de datos con los que es posible identificar el material, dado que estos datos están representados por números complejos, donde estos números serán extraídos de las lecturas, separados en módulo y fase para poder realizar cálculos correctos.

Cada extracción se ha realizado mediante la herramienta facilitada por Acconeer.

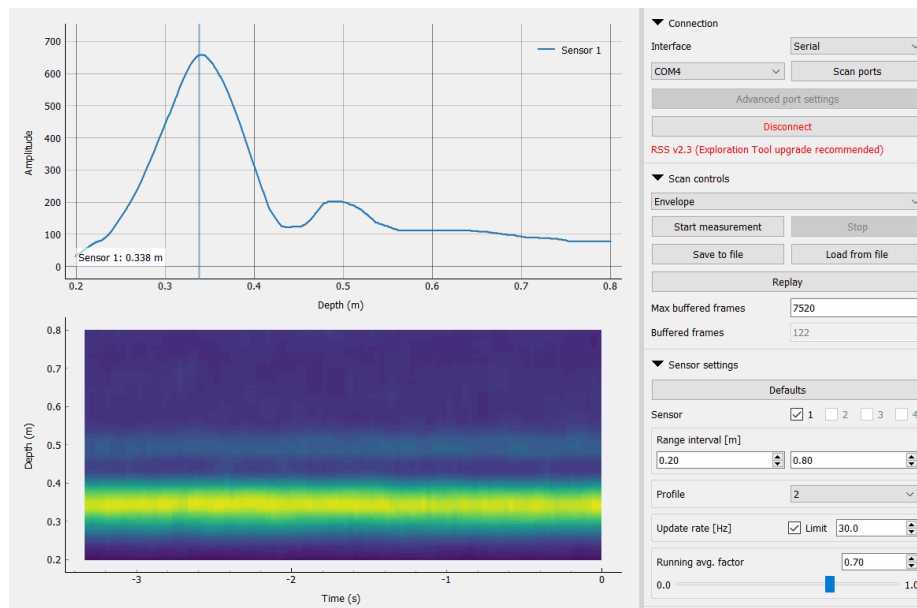


Figura 3.5: Herramienta Acconeer

Se necesita la aplicación Acconeer Exploration Tool que se puede descargar desde el repositorio de este TFG https://github.com/mecyc/TFG_RADAR_60GHZ/tree/main/acconeer-python-exploration. En esta interfaz se conecta utilizando el modo socket a la IP de la Raspberry Pi 4.

Una vez dentro hay cuatro servicios y varias funcionalidades. Los servicios son:

- Power Bins: según <https://aconeer-python-exploration.readthedocs.io/en/latest/services/pb.html>, calcula la energía de diferentes distancias, su objetivo es la medición de objetos grandes a distancias cortas como un sensor de parking. Al ser un método muy sencillo con poca cantidad de datos y enfocado a grandes objetos como una pared no se ha contemplado para el estudio.
- Envelope: según <https://aconeer-python-exploration.readthedocs.io/en/latest/services/envelope.html>, es igual que Power Bins pero utilizando un espectro continuo de los datos. Su caso de uso típico es la detección estática, por esto se ha contemplado para el estudio ya que la basura estará quieta.
- IQ: según <https://aconeer-python-exploration.readthedocs.io/en/latest/services/iq.html>, utiliza la coherencia de fase del radar que detecta movimiento a nivel fino. Tiene cinco modos, el primero es el que ha sido relevante al detectar mejores reflejos de los datos y estar optimizado para distancias muy cortas.
- Sparse: según <https://aconeer-python-exploration.readthedocs.io/en/latest/services/sparse.html>, se basa en la señal está más cuantizada y su uso principal es el análisis del movimiento y no de situaciones estáticas, por ello no se ha contemplado su uso para el estudio.

Para la creación del modelo se han seleccionado 30 materiales divididos en, 10 de plástico, 10 de cristal y 10 de cartón. De cada material se han realizado 10 lecturas, de varias caras, girando el objeto.

Llegamos a una colección de 300 lecturas exportadas cada una en ficheros con formato numpy (.npy) donde están almacenadas las características en vectores y matrices. Un porcentaje de estos datos conformarán la red de entrenamiento y otro porcentaje servirán para testear la red.

Cada instante de tiempo de la lectura comprende 291 atributos, de cada fichero obtenemos del orden de 300 instancias. Una instancia son estadísticas (medias, stds, etc) de los 291 atributos calculadas a partir de los aproximadamente 300 instantes de tiempo.

Tenemos los siguientes datos:

- N° Experimentos: Material - vista
- M instantes de tiempo

- Atributos

Técnicas y herramientas

4.1. Técnicas y metodologías

Metodología SCRUM

Se trata de una metodología de trabajo ágil que tiene como finalidad dividir en periodos de tiempo el flujo de trabajo. Estos periodos son conocidos como *sprints*, al finalizar cada *sprint* se realizan revisiones y reuniones donde se deciden las tareas de los próximos *sprints*.

Cliente de control de versiones

- Herramientas consideradas: GitHub Desktop y Gitkraken
- Herramienta elegida: GitHub Desktop

GitHub Desktop utilizado para la gestión ágil del proyecto.

Hosting del repositorio

- Herramientas consideradas: GitLab y GitHub.
- Herramienta elegida: GitHub

GitHub ofrece una gran cantidad de facilidades para mantener el proyecto en la nube y debido a que no cobra por sus servicios lo convierte en la mejor opción posible. Nos permite alojar nuestro repositorio central del proyecto usando el control de versiones Git.

Git es un sistema de control de versiones distribuido de código abierto. El control de versiones nos permitirá retornar a algún punto anterior del desarrollo de nuestra aplicación en caso de sufrir errores.

Url de la herramienta: <https://github.com/>

4.2. Lenguajes y bibliotecas

Python

El lenguaje de programación Python se diferencia por su código el cual es legible y limpio esto hace que sea uno de los lenguajes de iniciación de muchos programadores. Además, se trata de un lenguaje de multiparadigma y multiplataforma muy utilizado en la técnica del BigData.

Url de la herramienta: <https://www.python.org/>

TensorFlow

TensorFlow es una biblioteca de software de código abierto para computación numérica. Desarrollada por Google capaz de construir y entrenar redes neuronales.

Tensorflow es capaz de realizar cálculos numéricos usando grafos de flujo de datos. Los nodos hacen referencia a las operaciones matemáticas y los enlaces representan conjuntos de datos multidimensionales que relacionan los nodos.

La arquitectura flexible de Tensorflow permite realizar los cálculos en una o varias CPUs o GPUs mediante una sola API.

Url de la herramienta: <https://www.tensorflow.org/>

NumPy

NumPy es una biblioteca utilizada en la programación con Python para crear vectores y matrices grandes multidimensionales.

La funcionalidad principal de NumPy es su estructura de datos "ndarray", para una matriz de n dimensiones.

El uso de NumPy en Python brinda una funcionalidad comparable a MATLAB.

Url de la herramienta: <https://numpy.org/>

Pandas

Pandas es una biblioteca de código abierto especializada en el manejo y análisis de estructuras de datos flexible y fácil de usar, construida sobre el lenguaje de programación Python.

Su objetivo es ser el bloque de construcción fundamental de alto nivel para realizar análisis de datos prácticos del mundo real en Python.

Pandas está construido sobre NumPy y está destinado a integrarse bien dentro de un entorno informático científico con muchas otras bibliotecas de terceros.

Url de la herramienta: <https://pandas.pydata.org/>

acconeer.exptool

acconeer.exptool es la biblioteca facilitada por la empresa Acconeer. Con ella pondremos en funcionamiento el radar de 60GHz además de recopilar los datos necesarios.

Url de la herramienta: <https://github.com/acconeer/acconeer-python-exploration>

Sklearn

Scikit-learn es una biblioteca para aprendizaje automático de software libre para el lenguaje de programación de Python. En esta biblioteca encontramos varios algoritmos de clasificación, regresión y análisis de grupos entre los cuales están máquinas de vectores de soporte, bosques aleatorios, Gradient boosting, K-means y el utilizado en este proyecto RandomForest.

Url de la herramienta: <https://scikit-learn.org/>

Tkinter

Tkinter es una adaptación de la biblioteca gráfica Tcl/Tk [21] para el lenguaje de programación Python. Mediante Tkinter se desarrollará la interfaz gráfica para comparar distintos tipos de materiales.

Url de la herramienta: <https://docs.python.org/es/3/library/tkinter.html>

4.3. Herramientas de desarrollo

Jupyter Notebook

Jupyter Notebook es el entorno de trabajo utilizado en el proyecto que permite desarrollar código en Python de manera dinámica. Nos ofrece integrar en un mismo archivo bloques de código, texto, gráficas o imágenes. Utilizado ampliamente en análisis numéricos y estadísticos.

Jupyter admite más de 40 lenguajes de programación, incluidos Python, R, Julia y Scala.

Url de la herramienta: <https://jupyter.org/>

4.4. Herramientas de documentación

\LaTeX

\LaTeX es un sistema de elaboración de documentos basado en comandos. El proyecto está desarrollado mediante una plantilla creada en \LaTeX .

Url de la herramienta: <https://jupyter.org/>

MiKTeX

MiKTeX es una distribución de \LaTeX encargada de gestionar los componentes y paquetes. Tiene la capacidad de actualizarse así mismo descargando nuevas versiones de componentes.

Url de la herramienta: <https://miktex.org/>

Texmaker

Texmaker es el editor de documentos \LaTeX utilizado para funcionar necesita MiKTeX.

Url de la herramienta: <https://www.xm1math.net/texmaker/>

BibItNow!

BibItNow! es la extensión del navegador que utilizamos para citar páginas web en el formato Bibtex, formato soportado por \LaTeX .

Url de la herramienta: <https://chrome.google.com/webstore/detail/bibitnow/bmnfikjlonhkoojjfddnlbinkkapmldg>

Aspectos relevantes del desarrollo del proyecto

En este apartado se va a recoger el ciclo de vida del proyecto, detallando los aspectos más relevantes que se han tratado y como se han resuelto las diferentes dificultades encontradas a lo largo de su desarrollo. Se irán presentando diferentes secciones que concuerdan con el orden cronológico seguido en el proyecto y muestran la justificación de las decisiones tomadas.

5.1. Propuesta del Proyecto

La propuesta de este proyecto consistía en crear un modelo clasificador integrado en una interfaz para el reconocimiento de materiales que se dividía en las siguientes 3 tareas principales:

- Montaje del sensor: Puesta en funcionamiento del radar junto con sus componentes. Siendo capaces de extraer las características de las lecturas realizadas.
- Lecturas de los materiales: Conseguir información a través de las características de los materiales para crear un modelo capaz de diferenciar materiales como son el cartón, el plástico y el cristal.
- Clasificador de materiales: La tarea principal pedida para realizar este proyecto era la de construir un clasificador de materiales a través de un radar de 60GHz devolviendo la identificación correcta del material.

5.2. Metodologías aplicadas

En el desarrollo de este proyecto se ha intentado emplear en la medida de lo posible la metodología ágil de Scrum.

Debido a que el proyecto se comenzó en abril de 2021 en la modalidad presencial, no se ha podido seguir de forma estricta todas las pautas de la metodología, como las reuniones diarias con todo el equipo.

Las pautas que se han seguido durante el proyecto han sido:

- Desarrollo incremental del proyecto mediante sprints.
- Sprints de 1 a 2 semanas.
- Al finalizar el sprint, reuniones para evaluar el proyecto y plantear los pasos a seguir en el siguiente sprint.

5.3. Formación

Durante las primeras fases del proyecto, fue necesario aprender el funcionamiento del Radar Sensor. Para esto, se consultó la documentación facilitada por Acconeer.

A parte de esto se investigó información adicional sobre el uso de estos radares, así como ejemplos de su uso, la información extraída de la red fue escasa o casi nula. únicamente fue válida la documentación facilitada por Acconeer² en el repositorio de GitHub.

5.4. Montaje del radar

En esta sección se indican tanto las necesidades hardware como software para el montaje y puesta en funcionamiento del radar.

Necesidades hardware

Los creadores de Acconner crearon un video ([EVK 2](#)) con las instrucciones del montaje del sensor. En él se explica el montaje de los diferentes componentes a la Raspberry Pi 4.

En un comienzo será necesario disponer de:

²<https://acconeer-python-exploration.readthedocs.io/>

- Raspberry Pi 4
- Radar A111
- Placa XR112
- Cable flexible para XR112
- Tarjeta SD
- Teclado USB
- Ratón USB
- Monitor con HDMI
- Cable HDMI

FOTO

La placa se conecta a la Raspberry por el puerto GPIO quedando por encima del resto.

FOTO

Se ha realizado un espacio o cubículo de lectura de objetos de $30 \times 22 \times 25 \text{ cm}$ (ancho, largo y alto) con una incisión en la parte inferior en el medio del tamaño del sensor. El sensor fue bloqueado a su posición solamente colocándolo en la posición de la incisión. Esta caja de madera se utiliza poniendo la abertura en la parte inferior, mirando al suelo.

El sensor tiene sensibilidad a partir de los 10 cms. La sensibilidad del sensor se limitó al rango de 10 a 24 cm, el mínimo de sensibilidad hasta el suelo.

FOTOPROTOTIPO

Necesidades software

Una vez se había procedido con el montaje del dispositivo se pasó a la instalación y configuración del sistema operativo. Además, se instalaron los diversos programas necesarios tanto para recoger los datos como para poder interpretarlos.

Se necesitarán las siguientes aplicaciones para completar la instalación. Además, serán muy útiles a la hora de trabajar con el Radar. Hay que descargar e instalar:

- **Acconeer SW**: disponible para todos los usuarios (Windows, Linux, IOS)
- **Raspbian OS**
- **Etcher**: se usará para instalar el sistema operativo Raspbian en la SD
- **PuTTY**: se utiliza para conectarse a Raspberry Pi
- **WinSCP**: utilizado para transferir archivos a Raspberry Pi

5.5. Lectura de los objetos

Para la creación del modelo se han seleccionado 30 materiales divididos en, 10 de plástico, 10 de cristal y 10 de cartón. De cada material se han realizado 10 lecturas, de varias caras, girando el objeto.

Llegamos a una colección de 300 lecturas exportadas cada una en ficheros con formato numpy (.npz) donde están almacenadas las características en vectores y matrices. Un porcentaje de estos datos conformarán la red de entrenamiento y otro porcentaje servirán para testear la red.

Cada instante de tiempo de la lectura comprende 291 atributos, de cada fichero obtenemos del orden de 300 instancias. Una instancia son estadísticas (medias, stds, etc) de los 291 atributos calculadas a partir de los aproximadamente 300 instantes de tiempo.

Tenemos los siguientes datos:

- N° Experimentos: Material - vista
- M instantes de tiempo
- Atributos

Se creará la siguiente matriz:

$$Matriz1 = (N \cdot M) \cdot A \quad (5.1)$$

Es una matriz 2D ($N \cdot M, A$)

La extracción y creación de la matriz con los datos necesarios se realizará mediante Python:

```
diccionario = np.load('C01_V01.npy',allow_pickle=True).item()
data = diccionario.get('sweep_data').get('data')
data = data.reshape(data.shape[0],data.shape[2])
```

Los datos de la lectura tienen una parte real y otra imaginaria, a partir del array 2D anterior, se obtiene un array 2D, con el doble de anchura. Por ejemplo, si Matriz1 es de 300×291 , se obtendrá otra matriz que comprende el módulo y la fase de 600×291 . Así se obtiene el módulo del número complejo y la fase del número complejo. El módulo será una matriz de 300×291 y la fase también.

```
modulo = abs(data)
fase = []
for i in data:
    fila = []
    for j in i:
        fila.append(phase(j)) #Fase
    fase.append(fila)
    fase = np.asarray(fase)
    modulo_fase = np.concatenate((modulo, fase), axis=0)
```

A partir de los datos del módulo y la fase se obtiene la media.

```
traspuesta = np.transpose(modulo_fase)
#Obtenemos la matriz traspuesta(291x600) de modulo_fase(600x291)

#Separamos la traspuesta en modulo y fase
t_modulo = traspuesta[:,int(traspuesta.shape[1]/2)]
t_fase = traspuesta[:,int(traspuesta.shape[1]/2):]
media = []
for i in t_modulo:
```

```
media.append(i.mean())  
for j in t_fase:  
    media.append(j.mean())  
#media -> medias modulo y fase
```

5.6. Modelo clasificador

Para elegir el modelo a usar se tuvieron que tener en cuenta diferentes factores como el peso del modelo, su carga de computo y su precisión.

Se realizaron pruebas con los diferentes modelos (KNN, RamdonFores y Logistic Regression) y al final opté por usar el modelo RandomFores debido a que el promedio hace que este modelo sea mejor que un solo árbol de decisión, por lo que mejora su precisión y reduce el sobreajuste.

La Regresión Logística obtuvo un bajo porcentaje de coincidencia un 59,33 %

```
from sklearn.metrics import accuracy_score  
  
accuracy = accuracy_score(y_test2, y_prds)  
accuracy2 = accuracy_score(y_train, y_prds_train)  
  
accuracy, accuracy2  
  
(0.5933333333333334, 0.8851851851851852)
```

Figura 5.1: Regresión Logística

Trabajos relacionados

En este apartado se muestran las herramientas relacionadas con el proyecto. Hoy en día hay poca información y documentación de este tipo de tecnología de radares en la red, por ello solo muestro una única herramienta.

6.1. RadarCat

RadarCat (Radar Categorization for Input and Interaction) permite que un dispositivo electrónico pueda reconocer y clasificar distintos materiales y objetos a tiempo real con alta precisión. Además da mucha más información de los objetos, como su estructura interna.

El escaner/radar utilizado envía ondas electromagnéticas al objeto, estas rebotan y vuelven al punto de partida para su procesamiento con aprendizaje automático o machine learning.

La figura 6.1 muestra un esquema de cómo está estructurado RadarCar. Podemos ver que es necesario un ordenador para inicializar el software de RadarCat, a este equipo se conecta una placa base donde se encuentran las antenas/radares capaces de realizar lecturas de objetos. Por último encontramos una placa que protege el sensor y soporta el objeto a reconocer.

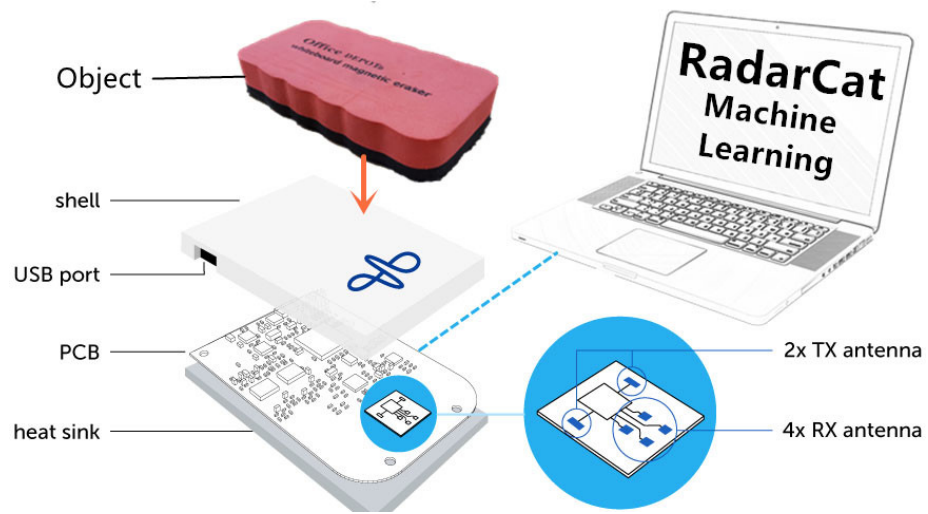


Figura 6.1: RadarCat.

Más allá de la interacción humana con la computadora, RadarCat también abre nuevas oportunidades en áreas como la navegación y el conocimiento mundial (por ejemplo, usuarios con baja visión).

Conclusiones y Líneas de trabajo futuras

7.1. Conclusiones

En este trabajo se exploraron las posibilidades de utilizar un radar de ondas de 60GHz para la clasificación de diferentes tipos de materiales. Se descubrió que extrayendo características mediante el radar y combinando esto con un clasificador Random Forest es posible realizar una clasificación para distinguir tres tipos de materiales comunes en la vida diaria como son PLÁSTICO, CRISTAL y CARTÓN con una precisión del 99,89 % usando validación cruzada.

Mediante una aplicación en fase Beta se ha conseguido realizar la lectura de las características de varios materiales y su correcta identificación.

Una posible aplicación sería automatizar la diferenciación de los tipos de embalajes en una empresa de envíos para su correcto reparto evitando los golpes.

Respecto a las dificultades encontradas, algunas han surgido por intentar abarcar una gran cantidad de lecturas de distintos objetos tanto en su estructura como en su forma, lo que ha significado tener un gran número de características dificultando la tarea de implementar un modelo clasificador que diese una gran exactitud a la hora de indicar el tipo de material.

7.2. Líneas de trabajo futuras

Posibles líneas de trabajo futuras relativas al uso del radar:

- Combinar el radar utilizado de Acconeer con un sensor radar del fabricante Infineon Technologies, posiblemente colocados en dos posiciones del habitáculo de lectura (un radar superior y otro lateral).
- Refactorización del código actual para su uso en otros sensores.
- Aplicar otros clasificadores.
- Añadir otros estadísticos.
- Automatizar el proceso de creación y entrenamiento del clasificador mediante una base de datos, así como la ampliación de la capacidad de reconocer materiales nuevos.

Bibliografía

- [1] J. L. Alonso Cerpa, “Introducción a los sistemas radar y arpa.” <https://riull.ull.es/xmlui/bitstream/915/6842/1/INTRODUCCION%20A%20LOS%20SISTEMAS%20RADAR%20Y%20ARPA.pdf>, 2015.
- [2] “Herrera, g., tomás, r., lópez sánchez, j. m., oriol, m., cooksley, g., & mulas de la peña, j. (2009). sistemas radar aplicados a la investigación de subsidencia y movimientos de ladera. enseñanza de las ciencias de la tierra, 17(3), 316–324. - buscar con google.” [https://www.google.com/search?q=Herrera%2C+G.%2C+Tom%C3%A1s%2C+R.%2C+L%C3%B3pez+S%C3%A1nchez%2C+J.+M.%2C+Oriol%2C+M.%2C+Cooksley%2C+G.%2C+%26+Mulas+de+la+Pe%C3%B1a%2C+J.+\(2009\).+Sistemas+radar+aplicados+a+la+investigaci%C3%B3n+de+subsidencia+y+movimientos+de+ladera.+Ense%C3%Bianza+de+Las+Ciencias+de+La+Tierra%2C+17\(3\)%2C+316%E2%80%93324.&aq=chrome..69i57.276j0j4&sourceid=chrome&ie=UTF-8](https://www.google.com/search?q=Herrera%2C+G.%2C+Tom%C3%A1s%2C+R.%2C+L%C3%B3pez+S%C3%A1nchez%2C+J.+M.%2C+Oriol%2C+M.%2C+Cooksley%2C+G.%2C+%26+Mulas+de+la+Pe%C3%B1a%2C+J.+(2009).+Sistemas+radar+aplicados+a+la+investigaci%C3%B3n+de+subsidencia+y+movimientos+de+ladera.+Ense%C3%Bianza+de+Las+Ciencias+de+La+Tierra%2C+17(3)%2C+316%E2%80%93324.&aq=chrome..69i57.276j0j4&sourceid=chrome&ie=UTF-8), 2009.
- [3] “Rigo ribas, t. (2004). estudio de sistemas convectivos mesoscales en la zona mediterránea occidental mediante el uso del radar meteorológico. [http://cataleg.ub.edu/record=b1691394~s1 %20acat](http://cataleg.ub.edu/record=b1691394~s1%20acat) - buscar con google.” [https://www.google.com/search?q=Rigo+Ribas%2C+T.+\(2004\).+Estudio+de+sistemas+convectivos+mesoscales+](https://www.google.com/search?q=Rigo+Ribas%2C+T.+(2004).+Estudio+de+sistemas+convectivos+mesoscales+)

- en+la+zona+mediterr%C3%A1nea+occidental+mediante+el+uso+del+radar+meteorol%C3%B3gico.+http%3A%2F%2Fcataleg.ub.edu%2Frecord%3Db1691394~S1%252Acat&oq=Rigo+Ribas%2C+T.+(2004).+Estudio+de+sistemas+convectivos+mesoscalares+en+la+zona+mediterr%C3%A1nea+occidental+mediante+el+uso+del+radar+meteorol%C3%B3gico.+http%3A%2F%2Fcataleg.ub.edu%2Frecord%3Db1691394~S1%252Acat&aqs=chrome..69i57.212j0j4&sourceid=chrome&ie=UTF-8, 2004.
- [4] M. Acosta Osorio, *Estudio de técnicas de procesamiento de señal en sistemas de radar para la detección del rango de un objeto/obstáculo para pruebas en ambientes de laboratorio*. PhD thesis, Universidad Pontificia Bolivariana, Mar 2014.
- [5] D. Marchionni and F. Cavayas, “La teledetección por radar como fuente de información litológica y estructural: Análisis espacial de imágenes SAR de,” *Geoacta*, vol. 39, no. 1, pp. 62–89, 2014.
- [6] “Sacristán romero, f. (2006a). la teledetección satelital y los sistemas de protección ambiental. aquatic, 5(24), 13–41. <https://doi.org/10.22518/16578953.701> - buscar con google.” [https://www.google.com/search?q=Sacrist%C3%A1n+Romero%2C+F.+\(2006a\).+La+Teledetecci%C3%B3n+satelital+y+los+sistemas+de+protecci%C3%B3n+ambiental.+Aquatic%2C+5\(24\)%2C+13%E2%80%9341.+https%3A%2F%2Fdoi.org%2F10.22518%2F16578953.701&oq=Sacrist%C3%A1n+Romero%2C+F.+\(2006a\).+La+Teledetecci%C3%B3n+satelital+y+los+sistemas+de+protecci%C3%B3n+ambiental.+Aquatic%2C+5\(24\)%2C+13%E2%80%9341.+https%3A%2F%2Fdoi.org%2F10.22518%2F16578953.701&aqs=chrome..69i57.191j0j4&sourceid=chrome&ie=UTF-8](https://www.google.com/search?q=Sacrist%C3%A1n+Romero%2C+F.+(2006a).+La+Teledetecci%C3%B3n+satelital+y+los+sistemas+de+protecci%C3%B3n+ambiental.+Aquatic%2C+5(24)%2C+13%E2%80%9341.+https%3A%2F%2Fdoi.org%2F10.22518%2F16578953.701&oq=Sacrist%C3%A1n+Romero%2C+F.+(2006a).+La+Teledetecci%C3%B3n+satelital+y+los+sistemas+de+protecci%C3%B3n+ambiental.+Aquatic%2C+5(24)%2C+13%E2%80%9341.+https%3A%2F%2Fdoi.org%2F10.22518%2F16578953.701&aqs=chrome..69i57.191j0j4&sourceid=chrome&ie=UTF-8).
- [7] B. Martí Cardona, A. Ramos, E. Bladé i. Castellet, T. Duc Tran, and J. Dolz Ripollès, “La teledetección como fuente de información para los modelos bidimensionales: aplicación al modelo Íber de las marismas de Doñana,” 2011.
- [8] R. Acevo Herrera, *Sistemas de teledetección activos y pasivos embarcados en sistemas aéreos no tripulados para la monitorización de la tierra*. PhD thesis, Universitat Politècnica de Catalunya, Spain, Apr 2011.
- [9] F. S. Romero, “Teledetección satelital en la visión territorial y sistemas de protección ambiental urbano-rural (2ª parte),” *Urbano*, vol. 10, no. 15, pp. 84–90, 2007.

- [10] J. A. F. Roselló and L. M. Rego, “La teledetección en el manejo de riesgos de desastres. Sus impactos en la sociedad,” *stgo*, no. 118, pp. 76–88, 2009.
- [11] M. I. Botana and S. E. Fernández, “Teledetección como experiencia de aprendizaje : Una mirada desde Geografía Física I, Geografía de los Espacios Marítimos y Cartografía,” *Universidad Nacional de La Plata. Facultad de Humanidades y Ciencias de la Educación. Departamento de Geografía*, 2019.
- [12] “Radar sensor introduction — acconeer-python-exploration documentation,” Jan 2021. [Online; accessed 11. Jan. 2022].
- [13] M. A. Ré, G. G. A. Varela, and S. Masuelli, “SEGMENTACIÓN DE SECUENCIAS CON DISTRIBUCIÓN GAMMA,” *ANALES AFA*, vol. 30, pp. 36–41, Jul 2019.
- [14] C. M. . G. B. Cancio A., “[proceedings title].” https://ri.conicet.gov.ar/bitstream/handle/11336/45615/CONICET_Digital_Nro.4c434d36-67a6-4643-800c-c3fac5db90af_A.pdf?sequence=2&isAllowed=y, 2017.
- [15] C. F. Castillo Plasencia, *Diseño de módulos de generación, conversión de frecuencia, amplificación y sincronización para un radar perfilador de vientos que opera a 445 MHz*. PhD thesis, Pontificia Universidad Católica del Perú, Oct 2016.
- [16] “Sandoval serrano, l. (2018). algoritmos de aprendizaje automático para análisis y predicción de datos. revista tecnológica, 11, 36–40. - buscar con google.” <https://www.google.com/search?q=Sandoval+Serrano%2C+L.+%282018%29.+Algoritmos+de+aprendizaje+autom%C3%A1tico+para+an%C3%A1lisis+y+predicci%C3%B3n+de+datos.+Revista+Tecnol%C3%B3gica%2C+11%2C+36%E2%80%9340.&oeq=Sandoval+Serrano%2C+L.+%282018%29.+Algoritmos+de+aprendizaje+autom%C3%A1tico+para+an%C3%A1lisis+y+predicci%C3%B3n+de+datos.+Revista+Tecnol%C3%B3gica%2C+11%2C+36%E2%80%9340.&aqs=chrome..69i57j0j4&sourceid=chrome&ie=UTF-8>.
- [17] “Russo, c., ramón, h., alonso, n., cicerchia, b., esnaola, l., & tessore, j. p. (2016). tratamiento masivo de datos utilizando técnicas de machine learning. xviii workshop de investigadores en ciencias de la computación, 131–134. - buscar con google.” <https://www.google.com/search?q=Russo%2C+C.%2C+Ram%C3%B3n%2C+Alonso%2C+N.+%26+Cicerchia%2C+B.+%26+Esnaola%2C+L.+%26+Tessore%2C+J.+P.+%282016%29.+Tratamiento+masivo+de+datos+utilizando+t%C3%A9cnicas+de+machine+learning.+xviii+workshop+de+investigadores+en+ciencias+de+la+computaci%C3%B3n%2C+131%E2%80%93134.&oeq=Russo%2C+C.%2C+Ram%C3%B3n%2C+Alonso%2C+N.+%26+Cicerchia%2C+B.+%26+Esnaola%2C+L.+%26+Tessore%2C+J.+P.+%282016%29.+Tratamiento+masivo+de+datos+utilizando+t%C3%A9cnicas+de+machine+learning.+xviii+workshop+de+investigadores+en+ciencias+de+la+computaci%C3%B3n%2C+131%E2%80%93134.&aqs=chrome..69i57j0j4&sourceid=chrome&ie=UTF-8>.

- 2C+H.%2C+Alonso%2C+N.%2C+Cicerchia%2C+B.%2C+Esnaola%2C+L.%2C+%26+Tessore%2C+J.+P.+(2016).+Tratamiento+Masivo+de+Datos+Utilizando+T%C3%A9cnicas+de+Machine+Learning.+XVIII+Workshop+de+Investigadores+En+Ciencias+de+La+Computaci%C3%B3n%2C+131%E2%80%93134.&aq= Russo%2C+C.%2C+Ram%C3%B3n%2C+H.%2C+Alonso%2C+N.%2C+Cicerchia%2C+B.%2C+Esnaola%2C+L.%2C+%26+Tessore%2C+J.+P.+(2016).+Tratamiento+Masivo+de+Datos+Utilizando+T%C3%A9cnicas+de+Machine+Learning.+XVIII+Workshop+de+Investigadores+En+Ciencias+de+La+Computaci%C3%B3n%2C+131%E2%80%93134.&aqs=chrome..69i57.178j0j9&sourceid=chrome&ie=UTF-8, 2016.
- [18] G. S. Cornejo Macias, “Diseño de un modelo de aprendizaje automático (machine learning) para clasificar texto en variables pest..” <http://repositorio.ug.edu.ec/bitstream/redug/54980/1/TRABAJO%20TITULACI%C3%93N%20-%20CORNEJO%20MACIAS%20GEOMAIRA%20SELENA.pdf>, 2021.
- [19] F. González, “Modelos de aprendizaje computacional en reumatología tt - machine learning models in rheumatology.” <http://www.scielo.org.co/pdf/rcrc/v22n2/v22n2a01.pdf>, 2015.
- [20] C. J. Carmona, F. Pulgar-Rubio, A. M. García-Vico, P. González, and M. J. del Jesus, “Análisis descriptivo mediante aprendizaje supervisado basado en patrones emergentes,” *VII Simposio de Teoría y Aplicaciones de Minería de Datos*, 2015.
- [21] Colaboradores de los proyectos Wikimedia, “Tcl - Wikipedia, la enciclopedia libre,” Apr 2020. [Online; accessed 24. Jun. 2021].