

**IMPLEMENTACIÓN DE UN SISTEMA DE CLOUD COMPUTING BASADO EN
GOOGLE CLOUD PLATFORM PARA EL REGISTRO Y RECONOCIMIENTO DE
PERSONAS UTILIZANDO DEEP LEARNING**

NELSON ESTEBAN AYALA GÓMEZ

UNIVERSIDAD DE PAMPLONA
FACULTAD DE INGENIERÍAS Y ARQUITECTURA
PROGRAMA DE INGENIERÍA EN TELECOMUNICACIONES
PAMPLONA - COLOMBIA

2021

**IMPLEMENTACIÓN DE UN SISTEMA DE CLOUD COMPUTING BASADO EN
GOOGLE CLOUD PLATFORM PARA EL REGISTRO Y RECONOCIMIENTO DE
PERSONAS UTILIZANDO DEEP LEARNING**

NELSON ESTEBAN AYALA GÓMEZ

Director:

PhD. JESÚS EDUARDO ORTÍZ SANDOVAL

Co-Director:

Ing. GERMAN ARLEY PORTILLA

TRABAJO PARA OPTAR POR EL TÍTULO DE INGENIERO EN TELECOMUNICACIONES

UNIVERSIDAD DE PAMPLONA

FACULTAD DE INGENIERÍAS Y ARQUITECTURA

PROGRAMA DE INGENIERÍA EN TELECOMUNICACIONES

PAMPLONA - COLOMBIA

2021

DEDICATORIA

Primero le agradezco a Dios infinitamente, porque a pesar de la actual situación que estamos afrontando, siempre me dió la fuerza y decisión para continuar llegando hoy al final de mi segundo logro, mi segundo pregrado profesional.

A mi madre, Zoley Gómez Quintero, mi pilar, la razón más importante para seguir en este largo camino y poder retribuirle mil veces más todo el esfuerzo que realizó para tenerme hoy donde estoy, por confiar y creer siempre, porque cada logro mío, será aún más de ella.

A mi padre, Nelson Enrique Ayala, mi mayor ejemplo, mi modelo a seguir, mi cuna de sabiduría que a pesar de tenerlo lejos siempre lo sentí cerca durante estos años, gracias papá por todo y vamos por más!

A mis dos amigos fieles de 4 patitas que no lograron llegar al final pero que me acompañaron más de 13 años en todo mi proceso como profesional, ahora se que tengo dos angelitos peludos que siempre caminarán a mi lado por donde vaya.

AGRADECIMIENTOS

A toda la planta docente del programa de Ingeniería en Telecomunicaciones, los cuales a partir de sus experiencias personales y académicas me formaron ejemplarmente tanto como profesional y como persona para ser alguien capacitado y competente ante nuevos retos. Mas que docentes me llevo el recuerdo de ustedes como grandes amigos.

A mi director de trabajo de grado, Ing. Jesús Ortíz, el cual además de ser mi docente ha sido mi amigo incondicional durante tantos años, fueron tantas vivencias y experiencias en estos años que sé que no quedarán ahí. Éxitos en sus próximos proyectos a realizar y confío plenamente que llegará aún mas lejos de lo que ha llegado actualmente.

A una persona que nos conocimos como una linda casualidad y ahora se ha vuelto parte de mi vida, Jessica Gómez, este triunfo es de los dos y sé que Dios nos tiene para grandes cosas.

También quiero agradecer a mis compañeros/as de la universidad, fueron pocos con los que compartí durante este pregrado pero me llevo el recuerdo de su calidad humana y profesional, esas largas conversaciones que a pesar de que nos tocó modalidad virtual un tiempo, logramos hacer de ese tiempo mucho mas ameno, espero volverlos a ver en un futuro.

Gracias a todos.

RESUMEN

Este trabajo plantea el desarrollo e implementación de un sistema de registro y reconocimiento de personas que permite la detección de usuarios sin contar con la ayuda de terceros.

La metodología a utilizar tiene dos fases fundamentales: una etapa de creación del frontend y una de la realización del backend. En la etapa de la creación del frontend se realizará una aplicación móvil creada tanto para dispositivos Android como para IOS mediante Flutter, desde la cual el administrador podrá controlar las acciones a realizar. Toda la implementación backend del proyecto se trabajará como computación en la nube (Cloud Computing). El algoritmo de detección de rostros se realizará de dos maneras, la primera haciendo uso de inteligencia artificial mediante redes neuronales convolucionales en cascada y la segunda mediante Open CV.

Con lo expuesto anteriormente, se tendrá una comparación de resultados la cual entregará datos para posterior análisis y respectivas conclusiones.

ABSTRACT

This work proposes the development and implementation of a system of registration and recognition of people that allows the detection of users without the help of third parties.

The methodology to be used has two main phases: a frontend creation stage and a backend realization stage. In the frontend creation stage, a mobile application will be created for both Android and IOS devices using Flutter, from which the administrator will be able to control the actions to be performed. All the backend implementation of the project will work as Cloud Computing. The face detection algorithm will be done in two ways, the first using artificial intelligence through convolutional neural networks in cascade and the second using Open CV.

With the above, we will have a comparison of results which will provide data for further analysis and respective conclusions.

Índice general

1. INTRODUCCIÓN	14
1.1. Planteamiento del Problema y Justificación	14
1.2. Objetivos	17
1.2.1. Objetivo General:	17
1.2.2. Objetivos Especificos:	17
2. MARCO TEÓRICO Y ESTADO DEL ARTE	18
2.1. La Neurona	18
2.1.1. Función de entrada (input function)	19
2.1.2. Función de activación (activation function)	20
2.1.3. Función de salida (output function)	21
2.2. Redes Neuronales	22
2.2.1. Elementos básicos que componen una red neuronal	22
2.3. Deep Learning	23
2.3.1. Redes Neuronales Convolucionales	24
2.4. Machine Learning	27
2.4.1. Definición	27
2.4.2. Antecedentes	28

2.4.3.	Terminología básica y notaciones	29
2.4.4.	Tipos de Machine Learning	30
2.4.5.	Aprendizaje Supervisado	31
2.5.	Inteligencia Artificial	33
2.5.1.	Historia de la IA	33
2.5.2.	¿Qué es exactamente la inteligencia artificial?	35
2.5.3.	Ramas de la IA	36
2.5.4.	Aplicaciones	38
3.	METODOLOGÍA	42
3.1.	Normativas, leyes, reglamentos	43
3.2.	Backend	44
3.2.1.	Raspberry Pi 4 B	44
3.3.	Frontend	53
3.3.1.	Diseño de la aplicación móvil	53
4.	RESULTADOS	69
4.1.	Reconocimiento y acceso de datos	69
4.2.	Comparación entre los dos métodos de detección	76
4.3.	Estudio económico del proyecto	78
5.	CONCLUSIONES	80
6.	Referencias Bibliograficas	83

Índice de figuras

2.1. Neurona biológica. Estructura de una neurona humana identificando sus diferentes partes.	19
2.2. Modelo de una neurona artificial con diversas entradas, una función de activación y su salida.	19
2.3. Activación lineal. Función más utilizada al momento de activar una neurona artificial.	21
2.4. Activación sigmoidea. Función usada al momento de activar cuya base son dos curvas en el eje x tanto positivo como negativo.	21
2.5. Estructura de una red neuronal. Se muestra de forma clara las 3 capas contenidas en ella.	22
2.6. Diagrama de bloques del aprendizaje profundo evidenciando las dos etapas: entrenamiento y predicción.	24
2.7. Esquema completo de una red neuronal convolucional identificando la transición de las diversas etapas que suceden en ella y su posterior predicción final.	25
2.8. Proceso de convolución realizado en una imagen por el filtro programado y su respectiva capa de salida.	26
2.9. Operación de agrupación desde una capa 4x4, con su respectiva salida 2x2.	26

2.10. Proceso conjunto de convolucion con agrupamiento en redes neuronales convolucionales identificando con diferentes colores su compresión.	27
2.11. Machine Learning y los campos donde ha estado ligado desde su auge teniendo gran impacto.	28
2.12. Ejemplo de notación en ML para identificar claramente las diferencias entre características y etiquetas.	30
2.13. Esquema explicativo de la división del ML en sus 2 tipos y a su vez en sus 4 subtipos.	30
2.14. Modelo detallado de aprendizaje supervisado con características necesarias para su buen entrenamiento.	31
2.15. Tipos de clasificación usados en el ML para separar diferentes categorías (2 ó más).	32
2.16. Ejemplo de una regresión para crear una categoría de elementos que no están etiquetados.	33
2.17. Estatua del matemático Alan Turing en el parque Bletchley, Londres.	34
2.18. Ejemplos de áreas en donde la Inteligencia Artificial se puede usar con grandes resultados.	36
2.19. Las diferentes ramas de la Inteligencia Artificial en las cuales se puede profundizar y desarrollar grandes proyectos.	38
2.20. La visión artificial siendo usada en procesos médicos para mejores diagnósticos y posteriores resultados.	41
3.1. Arquitectura general del proyecto evidenciando claramente las 4 subdivisiones importantes(Backend, Frontend, Reconocimiento facial y App final).	42
3.2. Imagen de una Raspberry Pi 4B 8 GB desde la capa superior donde se pueden ver algunos de sus componentes.	45

3.3. Raspberry conectado con perifericos siendo usada como ordenador de bajo coste y aceptables prestaciones.	46
3.4. Raspberry usada en visión artificial conectada a servomotores para aplicaciones de reconocimiento de objetos o facial.	46
3.5. Comparación de varias características y parametros importantes entre modelos de Raspberry Pi.	47
3.6. Conexion porel protocolo SSH entre la raspberry Pi usada y el pc mediante el software Putty.	48
3.7. Vista lateral de módulo de Raspberry Pi Camera Board v1.3.Módulo utilizado en este proyecto.	50
3.8. Logo Firebase Authentication. Backend utilizado para el login de los usuarios de la app creada.	51
3.9. Modos de entrada que permite Firebase Auth para ingresar a la aplicacion (Facebook, Google, Twitter, entre otras.)	52
3.10. Esquema del funcionamiento de Firebase Authentication entre el servidor y el móvil.	53
3.11. Interfáz gráfica de Visual Studio Code, donde se trabaja el framework Flutter y su lenguaje nativo es Dart.	54
3.12. Pantallas con los escudos de la Universidad de Pamplona, del programa Ingenieria en Telecomunicaciones y del logo de la app para crear un splash de inicio.	55
3.13. Pantalla de logueo de la app con sus respectivos campos de usuario y contraseña.	56
3.14. Mensaje de alerta cuando el usuario ingresado no corresponde a alguno que se encuentra en la base de datos.	57
3.15. A la izquierda, Mensaje de error cuando la contraseña contiene menos de 6 caracteres. A la derecha, error de Contraseña invalida para el usuario.	58

3.16. Pantalla para ingreso de usuario el cual quisiera hacer reset a su contraseña olvidada.	59
3.17. A la izquierda, Notificacion de llegada de e-mail de la app para reestablecer contraseña. A la derecha, contenido del correo para reestablecer dicha contraseña.	60
3.18. Pantalla luego de dar clic al correo con su campo para introducir la nueva contraseña y poder guardarla.	60
3.19. Vinculo caducado al querer recobrar la contraseña ya sea por uso anterior o por tiempo agotado para usarlo.	61
3.20. Pantalla principal de la app: el home. En ella se encuentra el menú para ingresar a todas las opciones que permiten interactuar de manera correcta.	62
3.21. Menú lateral o drawer donde se observa el panel de notificaciones y el boton para cerrar la sesion y salir.	63
3.22. Primer momento en la pantalla de voz cuando no esta activo el listening por lo tanto no va a tomar ninguna orden.	64
3.23. Segundo momento en la pantalla de voz cuando se puede dar ordenes por voz y al app traducirá esta voz a texto.	65
3.24. Ejemplo de detección de rostro en la app móvil con el usuario desarrollador mostrando el buen funcionamiento.	66
3.25. Pantalla de registro de nuevos usuarios, los cuales seran los que el usuario root no conozca para que luego la app los reconozca mediante la cámara.	67
3.26. Ejemplo de Base de datos de un usuario en Cloud Firestore con sus respecitos datos (Datos no reales solo para uso del ejemplo).	68
4.1. Detección de rostro en la app móvil del usuario1(Esteban Ayala).	70

4.2. Datos guardados en firestore del usuario1 eliminando parte de ellos por su confi-	
dencialidad.	70
4.3. Detección de rostro en la app móvil del usuario2 (Zoley Gómez).	71
4.4. Datos guardados en firestore del usuario2.	71
4.5. Detección de rostro en la app móvil del usuario3 (Jessica Gomez).	72
4.6. Datos guardados en firestore del usuario3.	72
4.7. Detección de rostro en la app móvil del usuario4 (Juliana Ramirez).	73
4.8. Datos guardados en firestore del usuario4.	73
4.9. Detección de rostro en la app móvil del usuario5 (Carlos Carvajal).	74
4.10. Datos guardados en firestore del usuario5.	75
4.11. Detección de rostro en la app móvil del usuario6 (Jesús Ortiz).	75
4.12. Datos guardados en firestore del usuario6.	76

Capítulo 1

INTRODUCCIÓN

1.1. Planteamiento del Problema y Justificación

A lo largo de la historia se ha especulado con la posibilidad de crear sistemas artificiales capaces de pensar y actuar como seres humanos. Era de esperarse que una idea tan atractiva como la Inteligencia Artificial fuera obteniendo relevancia en esta era en la que la informática y la digitalización crecen exponencialmente. Vivimos en una sociedad que se caracteriza por la abundancia y movimiento de datos, tanto personales como empresariales, generados a través de los medios sociales, el Internet de las Cosas, la navegación online o la información generada como consumidores en las diferentes marcas. El tratamiento de toda esta cantidad de datos se hace imposible si no es a través de máquinas, siendo la Inteligencia Artificial un valor fundamental para poder extraer conclusiones y hacer de estos datos, conclusiones válidas para la toma de decisiones. Encontramos diferentes aplicaciones en donde se ha hecho uso de la IA: el data drive marketing, automatización de procesos, redes neuronales, y el objeto de nuestro estudio: los sistemas de seguridad y reconocimiento de personas. (Lastra,2018)

Estos sistemas se han implementado con diversas funciones, entre las más comunes se encuentran: el reconocimiento en los dispositivos móviles, el cual se usa tanto para desbloquear el móvil como en el área de multimedia organizando fotos y videos en subcarpetas; también para permitir o denegar el acceso a los trabajadores de una empresa; los bancos lo utilizan para ayudar a proteger a los consumidores y prevenir fraudes; detección del estado de ánimo de las personas.

Uno de los usos con más proyección a futuro de los sistemas de identificación y registro de personas es la seguridad en las viviendas, edificios, locales comerciales, etc. Algunos países han comenzado a utilizar estos sistemas para automatizar y brindar seguridad a la hora de vender productos en sus locales comerciales. El hecho de poder realizar compras en un supermercado sin la necesidad de tener un cajero o persona supervisando es un hito en el país al cual hay que apuntarle con estos sistemas de identificación.

Los sistemas de seguridad están pensados para casas de grandes dimensiones ya que usualmente estas son la que necesitan más de estos, sin embargo, estos pueden ser instalados en edificaciones de cualquier tamaño y lo único que promete es dar protección total. Según la corporación Horus Smart Control, un sistema domótico es mucho más confiable porque no se deja manipular tan fácilmente como otros sistemas, lo que hace que su índice de seguridad sea mayor. El ahorro de energía y el confort son las otras grandes ventajas de estos sistemas.

Al introducir estos sistemas en un establecimiento, vivienda o lugar de residencia, vamos haciendo de este un lugar automatizado, capaz de interactuar con un administrador y con diferentes

usuarios para tomar algunas decisiones y automatizar procesos de nuestra vida diaria.

La domótica es un término que se refiere a las técnicas que se utilizan para automatizar una vivienda. El tema de la automatización del hogar a nivel mundial se ha pensado desde los años 80 y su implementación arrancó en los años 90 en japon. La domótica en Colombia empezó a ingresar al país hace más de 10 años y ha evolucionado hasta el punto de tener empresas que crean su propio sistema de integración. (Junstrand Vasquez,2004)

En Colombia en la última década se ha visto el crecimiento de nuevas edificaciones, siendo estas cada vez más modernas tanto en su diseño como en sus sistemas eléctricos y demás. Es aquí donde surge la urgencia de domotizar estas edificaciones aportando servicios de gestión, seguridad y comunicación para estar a la vanguardia del mundo.

Por lo dicho anteriormente, se plantea el diseño y desarrollo de un sistema eficaz y económico totalmente basado en Cloud Computing que permita ser un asistente para personas desde niveles medios en uso de tecnología hasta niveles avanzados, el cual se proyecta como una iniciativa de “Startup” tecnológica, que, a diferencia de una Pyme, puede ser un negocio escalable más rápida y fácilmente.

1.2. Objetivos

1.2.1. Objetivo General:

Implementar un sistema de Cloud computing basado en Google CLOUD Platform para el registro y reconocimiento de personas utilizando Deep Learning.

1.2.2. Objetivos Especificos:

- Estructurar una arquitectura de sistema, desplegando microservicios de comunicación entre Backend (CC) y Frontend (Aplicación Móvil).
- Diseñar un sistema de reconocimiento de voz dual (Speech to text, text to Speech) utilizando modelos de inferencia desplegados en un sistema de CC.
- Desarrollar un sistema de clasificación facial de personas, a través del entrenamiento y validación de diversos tipos de arquitectura neuronales profundas.
- Realizar la recolección de los datos de identificación de las personas cumpliendo a cabalidad los principios estipulados de Habeas Data Colombia.
- Validar la sensibilidad del sistema relacionada con la detección de órdenes, clasificación de personas, y re-entrenamiento de nuevas características faciales.

Capítulo 2

MARCO TEÓRICO Y ESTADO DEL ARTE

2.1. La Neurona

Las células llamadas neuronas son una unidad de procesamiento que recibe un estímulo eléctrico de otras neuronas principalmente a través de su árbol dendrítico. El estímulo eléctrico recibido al pasar de un cierto umbral causa que la neurona a su vez imprima una señal eléctrica a través de su axón a otras neuronas.

Desde el punto de vista funcional, las neuronas constituyen procesadores de información sencillos. Poseen un canal de entrada de información (las dendritas), un órgano de cómputo (el soma), y un canal de salida (el axón). (Matich, 2001)

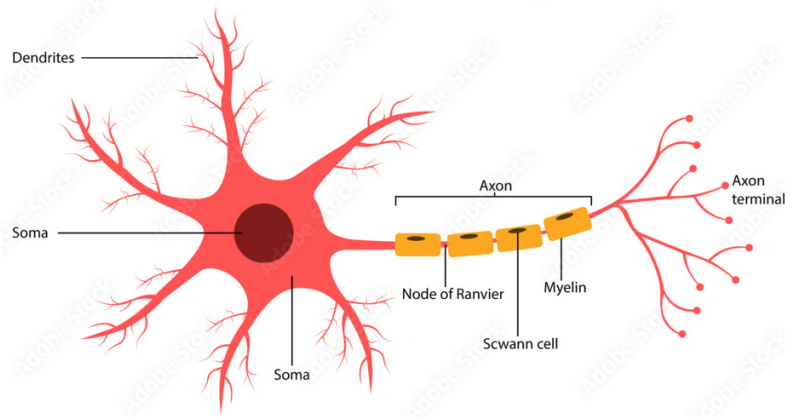


Figura 2.1: Neurona biológica. Estructura de una neurona humana identificando sus diferentes partes.

Las Redes Neuronales surgen como un intento de desarrollar sistemas que compitan con las características del cerebro, para conseguir su sofisticada capacidad de procesamiento de información.

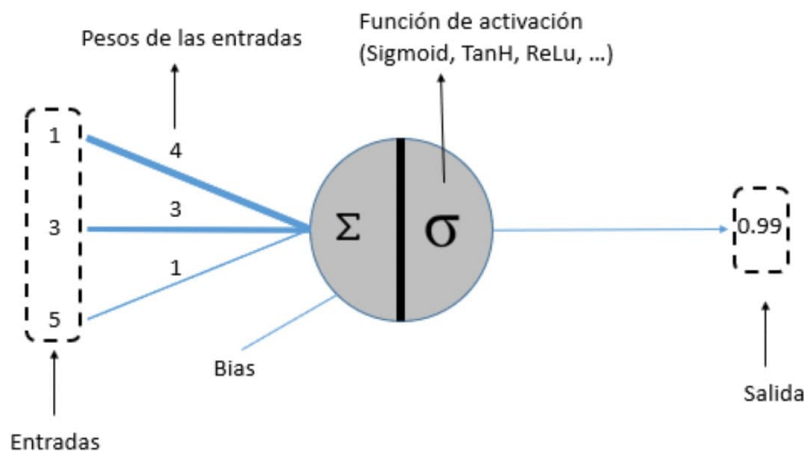


Figura 2.2: Modelo de una neurona artificial con diversas entradas, una función de activación y su salida.

2.1.1. Función de entrada (input function)

La neurona trata a muchos valores de entrada como si fueran uno solo; esto recibe el nombre de entrada global. Por lo tanto, ahora nos enfrentamos al problema de cómo se pueden combinar estas simples entradas (ini_1, ini_2, \dots) dentro de la entrada global, gini. Esto se logra a través de la función

de entrada, la cual se calcula a partir del vector entrada. La función de entrada puede describirse como sigue:

$$input_i = (in_i1w_i1) * (in_i2w_i2) * \dots (in_inw_in)$$

donde: * representa al operador apropiado (por ejemplo: máximo, sumatoria, producto, etc.), n al número de entradas a la neurona N_i y W_i al peso. Los valores de entrada se multiplican por los pesos anteriormente ingresados a la neurona. Por consiguiente, los pesos que generalmente no están restringidos cambian la medida de influencia que tienen los valores de entrada.

Algunas de las funciones de entrada más comúnmente utilizadas y conocidas son:

- 1) Sumatoria de las entradas pesadas
- 2) Producto de las entradas pesadas
- 3) Máximo de las entradas pesadas

2.1.2. Función de activación (activation function)

Una neurona biológica puede estar activa (excitada) o inactiva (no excitada); es decir, que tiene un “estado de activación”. Las neuronas artificiales también tienen diferentes estados de activación. La función activación calcula el estado de actividad de una neurona; transformando la entrada global (menos el umbral, i) en un valor (estado) de activación, cuyo rango normalmente va de (0 a 1) o de (-1 a 1). Esto es así, porque una neurona puede estar totalmente inactiva (0 o -1) o activa (1).

Las funciones de activación más comúnmente utilizadas se detallan a continuación:

- Función Lineal

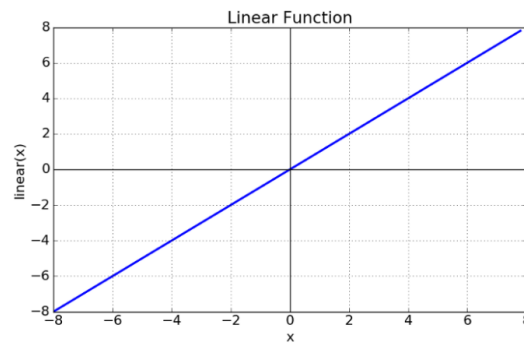


Figura 2.3: Activación lineal. Función más utilizada al momento de activar una neurona artificial.

- Función Sigmoidea

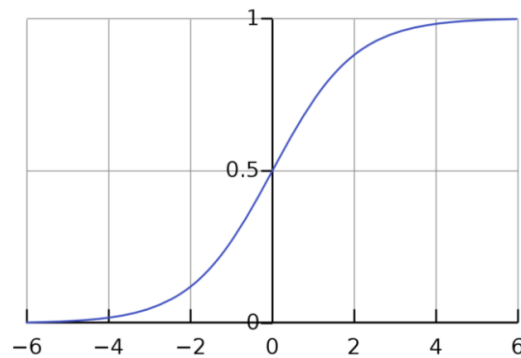


Figura 2.4: Activación sigmoidea. Función usada al momento de activar cuya base son dos curvas en el eje x tanto positivo como negativo.

2.1.3. Función de salida (output function)

El último componente que una neurona necesita es la función de salida. El valor resultante de esta función es la salida de la neurona i (out_i); por ende, la función de salida determina que valor se transfiere a las neuronas vinculadas. Si la función de activación está por debajo de un umbral determinado, ninguna salida se pasa a la neurona subsiguiente. (Izaurieta y Saavedra, 2000)

2.2. Redes Neuronales

2.2.1. Elementos básicos que componen una red neuronal

A continuación se puede ver, en la Figura 2.5, un esquema de una red neuronal:

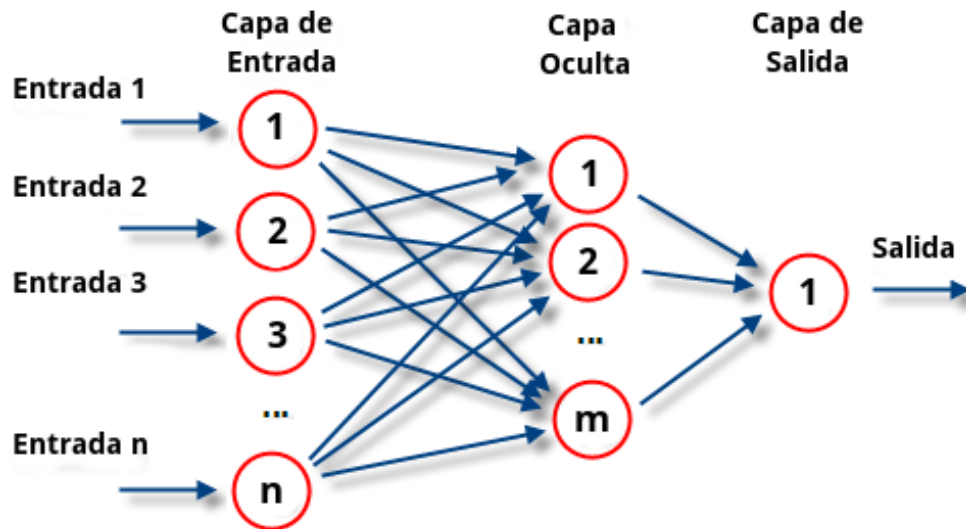


Figura 2.5: Estructura de una red neuronal. Se muestra de forma clara las 3 capas contenidas en ella.

La misma está constituida por neuronas interconectadas y arregladas en tres capas (esto último puede variar). Los datos ingresan por medio de la “capa de entrada”, pasan a través de la “capa oculta” y salen por la “capa de salida”. Cabe mencionar que la capa oculta puede estar constituida por varias capas.

La distribución de neuronas dentro de la red se realiza formando niveles o capas, con un número determinado de dichas neuronas en cada una de ellas. A partir de su situación dentro de la red, se pueden distinguir tres tipos de capas:

- De entrada: es la capa que recibe directamente la información proveniente de las fuentes externas de la red.
- Ocultas: son internas a la red y no tienen contacto directo con el entorno exterior. El número

de niveles ocultos puede estar entre cero y un número elevado. Las neuronas de las capas ocultas pueden estar interconectadas de distintas maneras, lo que determina, junto con su número, las distintas topologías de redes neuronales.

- De salidas: transfieren información de la red hacia el exterior

2.3. Deep Learning

Deep learning, conocido en español como aprendizaje profundo, son aquellas técnicas de aprendizaje automático que hacen uso de arquitecturas de redes neuronales. El concepto “profundo” viene referido al número de capas que poseen estas técnicas. Mientras que las redes neuronales convencionales poseen una capa, las redes neuronales profundas contienen varias capas. (Goodfellow, 2016)

El deep learning permite a las máquinas resolver problemas complejos incluso cuando se utiliza un conjunto de datos muy diverso, no estructurado e interconectado. Cuanto más aprenden los algoritmos de aprendizaje profundo, mejor se desempeñan.

Los algoritmos de aprendizaje profundo se apilan en una jerarquía de creciente complejidad y extracción. Cada algoritmo en la jerarquía aplica una transformación no lineal en su entrada y utiliza lo que aprende para crear un modelo estadístico como salida. Las iteraciones continúan hasta que la salida ha alcanzado un nivel de precisión aceptable (Yan y Geoffrey, 2015).

En este aprendizaje el modelo aprende por sí mismo y poder así descubrir las relaciones entre las variables. En los modelos basados en aprendizaje profundo no interviene el ser humano, el propio modelo es capaz de hacer la selección de características. En Deep Learning se hace uso de unas

redes neuronales especiales llamadas redes neuronales convolucionales. (Nuñez,2016)

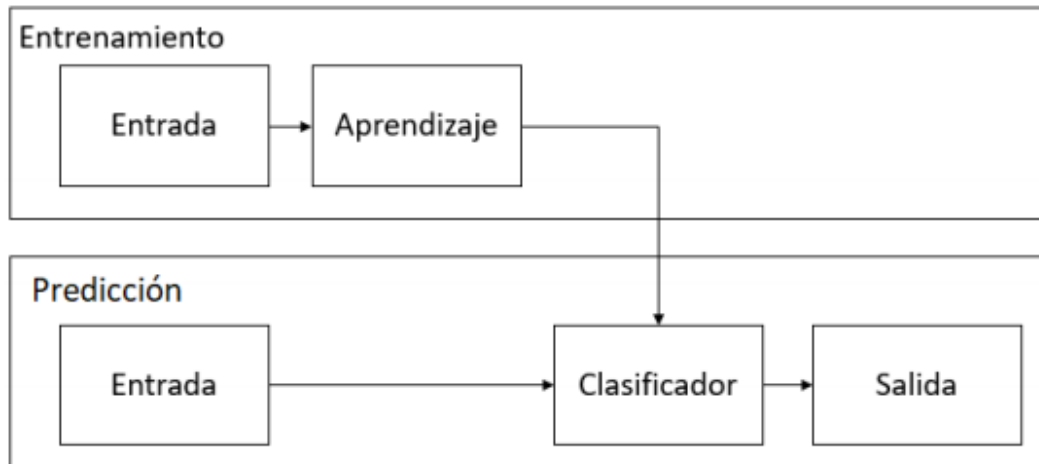


Figura 2.6: Diagrama de bloques del aprendizaje profundo evidenciando las dos etapas: entrenamiento y predicción.

2.3.1. Redes Neuronales Convolucionales

Estas redes son muy similares a las redes neuronales, están formadas por neuronas que tienen forma de pesos y sesgos que se pueden aprender. Están compuestas por tres capas pero existen unas capas ocultas que están especializadas en dos operaciones: convolución y agrupación.

La entrada de una red capa convolucional suele ser una imagen $m \times m \times r$, donde m es tanto la altura como el ancho de la imagen y r es el número de canales. Las capas convolucionales tienen k filtros (o kernels) cuyas dimensiones $n \times n \times q$, donde n y q son elegidas por el diseñador (generalmente q suele ser igual a r). Cada filtro genera mediante convolución un mapa de rasgos o características de tamaño $(m-n+1) \times (m-n+1) \times p$, siendo p el número de filtros que se desean usar. Después cada mapa es sub-muestreado en la capa de pooling con la operación “mean pooling” o “max pooling” sobre regiones contiguas de tamaño $p \times p$ donde p puede tomar valores desde 2 para imágenes pequeñas hasta, comúnmente, no más de 5 para imágenes grandes. Antes o después del submuestreo,

se aplica una función de activación sigmoïdal más un sesgo para cada mapa de rasgos. (Duran, 2017)

La composición de bloques anterior se puede ver en la Figura 2.7, en la que está representada una disposición genérica de las capas.

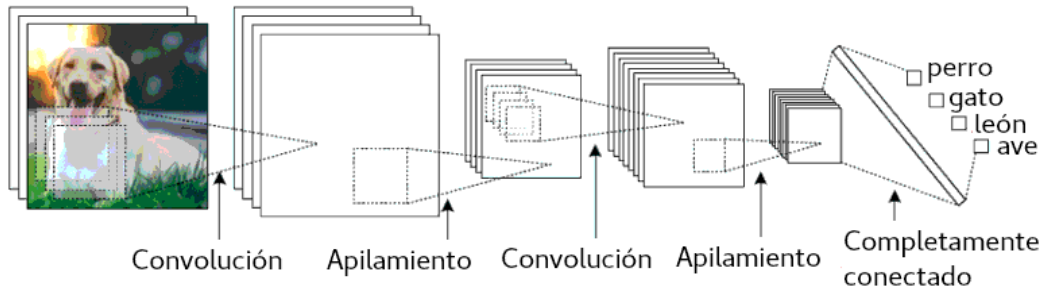


Figura 2.7: Esquema completo de una red neuronal convolucional identificando la transición de las diversas etapas que suceden en ella y su posterior predicción final.

Capa de Convolución

Estas capas ocultas se ocupan de aprender patrones locales en ventanas pequeñas de dos dimensiones, alto por ancho. Son capas que sirven para detectar características visuales en las imágenes como líneas, bordes y gotas de color. Estas capas permiten aprender una propiedad de la imagen en un punto específico y ser capaz de reconocerla en cualquier lugar de la propia imagen.

Las capas convolucionales son capaces de aprender diferentes elementos más complejos en función de lo aprendido en la capa anterior. Un ejemplo sería que una capa aprenda diferentes tipos de líneas que aparezcan en la imagen y la capa siguiente sea capaz de aprender elementos de la imagen que estén compuestos por diferentes líneas aprendidas en la capa anterior. La capacidad de realizar esto permite a estas redes poder aprender de una manera eficiente conceptos visuales que cada vez son más complejos. (Duran, 2017)

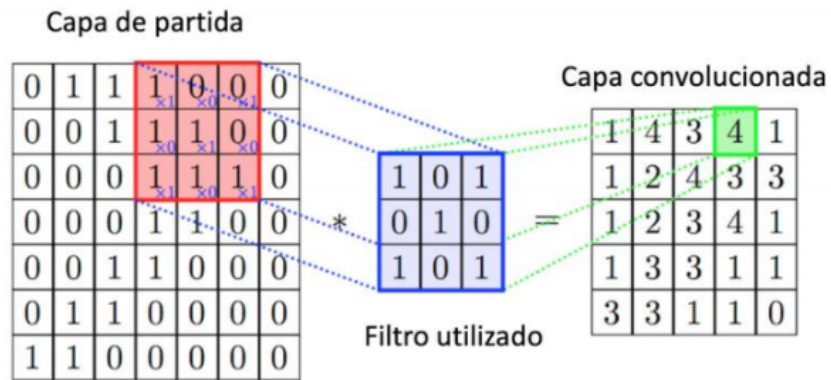


Figura 2.8: Proceso de convolución realizado en una imagen por el filtro programado y su respectiva capa de salida.

Como se puede observar en la Figura 2.8, esta matriz llamada Kernel recorre todas las neuronas de entrada. Por cada posición de la matriz hay una neurona. La matriz empieza a recorrer todas las neuronas por la parte superior izquierda y se desplaza hacia la derecha. Cuando ha acabado una fila, sigue con la fila de debajo. Así se consigue obtener información de todas las neuronas y que estén conectadas entre ellas.

Capa de Agrupación

A continuación de las capas convolucionales, las redes están acompañadas por unas capas de agrupación, también llamadas capas de reducción. Como su nombre indica, agrupa las neuronas en grupos.

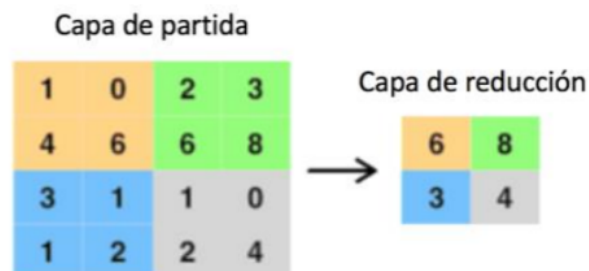


Figura 2.9: Operación de agrupación desde una capa 4x4, con su respectiva salida 2x2.

Al agrupar las neuronas la salida se ve reducida. Se reduce en función del tamaño de las agru-

paciones. Por ejemplo, si los grupos son de un tamaño 2x2, a la salida obtendremos la mitad del tamaño de los datos de entrada. En la figura 2.10 se muestra el proceso conjunto de las dos operaciones mencionadas previamente.(Duran,2017)

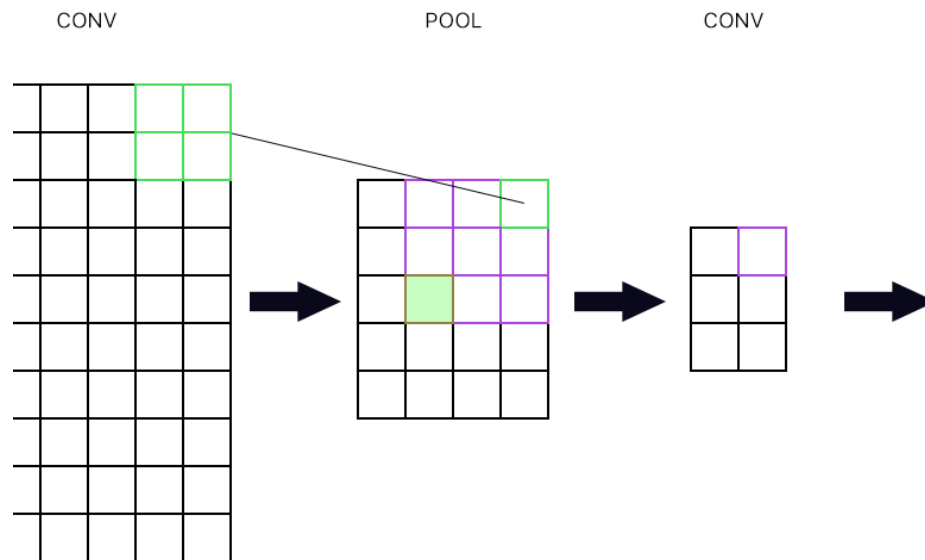


Figura 2.10: Proceso conjunto de convolucion con agrupamiento en redes neuronales convolucionales identificando con diferentes colores su compresión.

2.4. Machine Learning

2.4.1. Definición

Machine Learning es uno de los subcampos de la Inteligencia Artificial y puede ser definido como:

“Machine Learning es la ciencia que permite que las computadoras aprendan y actúen como lo hacen los humanos, mejorando su aprendizaje a lo largo del tiempo de una forma autónoma, alimentándolas con datos e información en forma de observaciones e interacciones con el mundo

real.” — Dan Fagella

Machine learning ofrece una manera eficiente de capturar el conocimiento mediante la información contenida en los datos, para mejorar de forma gradual el rendimiento de modelos predictivos y tomar decisiones basadas en dichos datos. (Román, 2019)

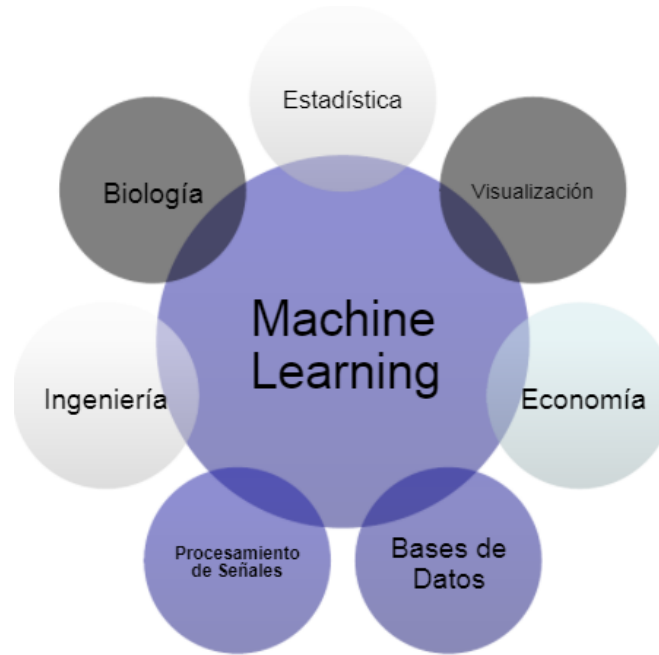


Figura 2.11: Machine Learning y los campos donde ha estado ligado desde su auge teniendo gran impacto.

2.4.2. Antecedentes

En 1959, el científico de la IBM Arthur Samuel escribió un programa para jugar damas, para mejorarlo hizo que el programa jugara consigo mismo miles de veces, el programa era capaz de mejorar su rendimiento a través de la experiencia, el programa aprendió y nació el Machine Learning.

A comienzos del 2000 este asombroso campo de la inteligencia artificial carecía de promoción

y recursos, esto debido a que desde sus orígenes estaba enfocada a sectores empresariales que tal vez no gozaban de suficientes ganancias como para alentar su pleno desarrollo, además en ese punto de la historia se desconocía el verdadero potencial del machine learning. (Jones, 2018)

No obstante, fue solo hasta los primeros meses del 2006 que el machine learning inició su gran ascenso a la cúspide de manejo y procesamiento de datos, ya que, de la mano de grandes empresas como IBM y Microsoft, empezaría a expandirse a nivel global. Luego tres años más tarde la industria IBM también revolucionaría el machine learning, lanzando su ordenador Watson el cual causó sensación en el concurso Jeopardy, un programa televisivo de norte américa que consistía en responder a preguntas en su lengua natural, pues este sistema venció a sus rivales humanos al responder con exactitud, aunque presentó problemas en algunas categorías que no le brindaban las pistas necesarias para que este sistema pudiera responder adecuadamente, aún así logró conseguir una excelente puntuación. (Hinestroza, 2018)

Al año siguiente, los amigos Jeff Dean, quien trabajaba en Google, y Andrew N, quien se desempeñaba como profesor de la universidad de Stanford, se asociaron para dirigir un innovador proyecto de Google que tendría como nombre Google Brain, que consistía en una red neuronal que por medio de Google es capaz de detectar patrones en videos e imágenes. (Hinestroza, 2018)

2.4.3. Terminología básica y notaciones

En Machine Learning generalmente se utilizan matrices y notaciones vectoriales para referirnos a los datos, de la siguiente forma:

- Cada fila de la matriz es una muestra, observación o dato puntual.

- Cada columna es una característica (o atributo), de la observación mencionada en el punto anterior.
- En el caso más general habrá una columna, que llamaremos objetivo, etiqueta o respuesta, y que será el valor que se pretende predecir.

Características					Etiqueta
Posición	Experiencia	Habilidad	País	Ciudad	Salario
Desarrollador	0	1	Colombia	Pamplona	2500000
Desarrollador	1	1	Colombia	Pamplona	2830000
Desarrollador	2	1	Colombia	Pamplona	2890000
Desarrollador	3	1	Colombia	Pamplona	3450000

Figura 2.12: Ejemplo de notación en ML para identificar claramente las diferencias entre características y etiquetas.

2.4.4. Tipos de Machine Learning

El Machine Learning tiene una amplia gama de aplicaciones, incluyendo motores de búsqueda, diagnósticos médicos, reconocimiento del habla y del lenguaje escrito, juegos y robótica. Pero para poder abordar cada uno de estos temas es crucial en primer lugar distinguir los distintos tipos de Machine Learning con los que nos podemos encontrar. (El Naqa, 2015)

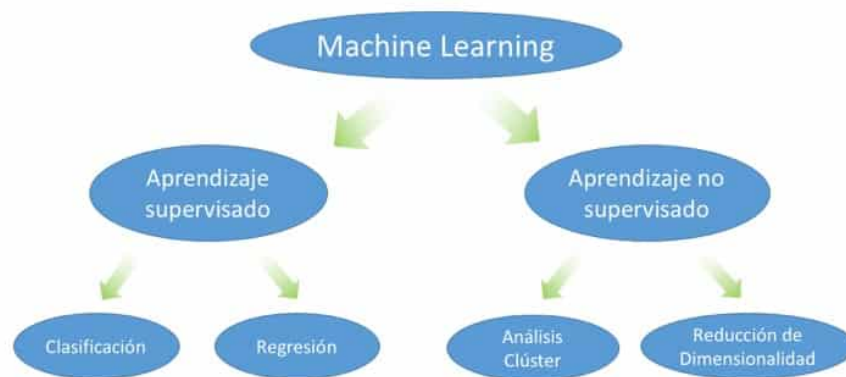


Figura 2.13: Esquema explicativo de la división del ML en sus 2 tipos y a su vez en sus 4 subtipos.

2.4.5. Aprendizaje Supervisado

En este tipo de aprendizaje, la máquina aprende de un conjunto de casos previamente etiquetados por un experto o de forma semi – automática basándose en los datos, es decir, para un conjunto de datos de entrada conocemos de antemano los datos correctos de salida. En el aprendizaje supervisado, el objetivo es aprender una forma de mapear las entradas en salidas, cuyos valores correctos son proporcionados por un supervisor. (Soto y Jimenez.2011)

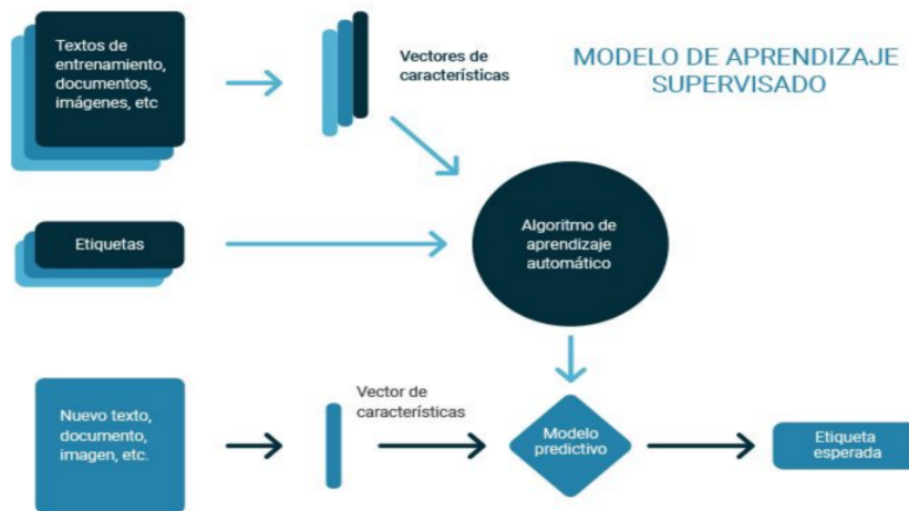


Figura 2.14: Modelo detallado de aprendizaje supervisado con características necesarias para su buen entrenamiento.

Clasificación

Clasificación es una sub-categoría de aprendizaje supervisado en la que el objetivo es predecir las clases categóricas (valores discretos, no ordenados, pertenencia a grupos). El ejemplo típico es la detección de correo spam, que es una clasificación binaria (un email es spam — valor “1”- o no lo es — valor “0” -). También hay clasificación multi-clase, como el reconocimiento de caracteres escritos a mano (donde las clases van de 0 a 9).

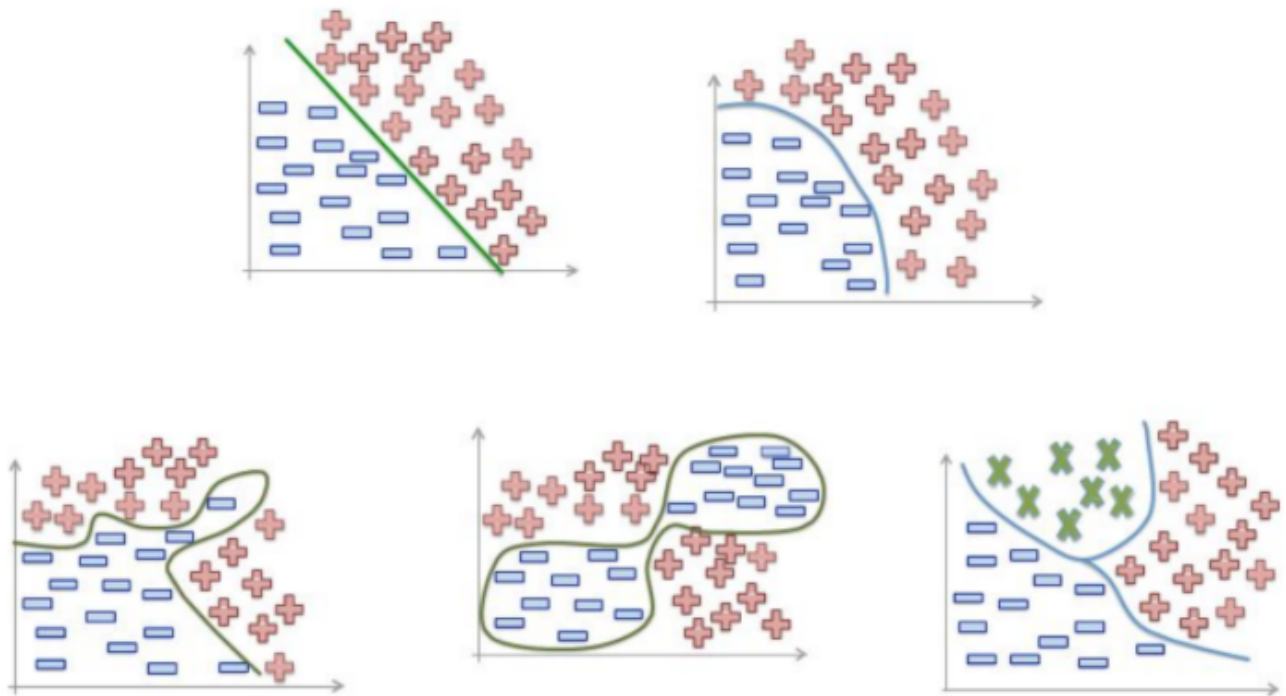


Figura 2.15: Tipos de clasificación usados en el ML para separar diferentes categorías (2 ó más).

Regresión

La regresión se utiliza para asignar categorías a datos sin etiquetar. En este tipo de aprendizaje tenemos un número de variables predictoras (explicativas) y una variable de respuesta continua (resultado), y se tratará de encontrar una relación entre dichas variables que nos proporciones un resultado continuo.

Un ejemplo de regresión lineal: dados X e Y, establecemos una línea recta que minimice la distancia (con el método de mínimos cuadrados) entre los puntos de muestra y la línea ajustada. Después, utilizaremos las desviaciones obtenidas en la formación de la línea para predecir nuevos datos de salida.

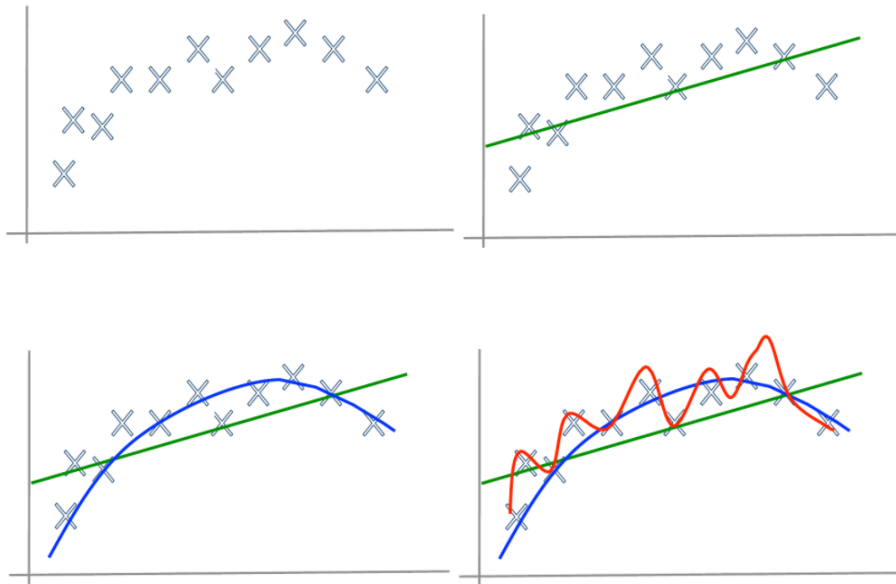


Figura 2.16: Ejemplo de una regresión para crear una categoría de elementos que no están etiquetados.

2.5. Inteligencia Artificial

2.5.1. Historia de la IA

La Inteligencia Artificial surge a partir de algunos trabajos publicados en la década de 1940 que no tuvieron gran repercusión, pero a partir del influyente trabajo de 1950 de Alan Turing, matemático británico, se abre una nueva disciplina de las ciencias de la información.

Se cree que los comienzos de la inteligencia artificial ocurrieron al definirse la neurona, en 1943, como un dispositivo binario con varias entradas y salidas. En el año de 1956 se volvió a tocar este tema y se establecieron las bases de la IA como un campo independiente dentro de la informática.

La IA fue introducida a la comunidad científica en 1950 por el inglés Alan Turing en su artículo "Maquinaria Computacional e Inteligencia." La pregunta básica que Turing trató de responder afirmativamente en su artículo era: ¿pueden las máquinas pensar? (Rich, Night, Calero, 1994)



Figura 2.17: Estatua del matemático Alan Turing en el parque Bletchley, Londres.

Dos de las contribuciones más importantes de Turing a la IA fueron el diseño de la primera computadora capaz de jugar ajedrez y, más importante que esto, el establecimiento de la naturaleza simbólica de la computación. Von Neumann fue el primero en antropomorfizar el lenguaje y la concepción de la computación al hablar de la "memoria", los "sensores", etc., de las computadoras. (Muggleton, 2014)

Desde fines de los cincuentas la investigación en IA se expande y se multiplica en direcciones diversas. La capacidad simbólica de las computadoras es estudiada, entre otros, por Shannon (1950) y por Newell, Shaw y Simon (1958) quienes diseñan el primer programa inteligente basado en su modelo de procesamiento de información. Algunos investigadores se dedicaron al estudio de la naturaleza del aprendizaje en las computadoras y a los procesos de reconocimiento de patrones visuales. Como resultado de ello Selfridge y Dinneen consiguen diseñar el primer programa capaz de aprender por experiencia (McCorduck, 1979).

A principios de los sesentas, la IA comienza una fase distinta de su desarrollo. En 1962, McCarthy y Raphael inician sus trabajos sobre el diseño y la construcción de un robot móvil que llamarían "Shakey". La idea básica de Raphael era la de reunir, en una sola, distintas máquinas con capacidad

de aprender por experiencia, de reconocer patrones visuales, de modelar, de manipular símbolos, etc., y esperar que el todo fuera mayor que la suma de las partes. Pero el avance tecnológico de la IA no se detuvo ahí, sino que continuó con otras áreas de estudio como Procesamiento de imágenes, que se trata de una clase de procesamiento de señales que puede tomar fotografías o fotogramas de vídeo y someterlos a un procesado de perfil bajo con el fin de extraer determinadas características o parámetros, o bien para elaborar nuevas imágenes procesadas como material de salida.

2.5.2. ¿Qué es exactamente la inteligencia artificial?

Se puede definir la IA como «la habilidad de los ordenadores para hacer actividades que normalmente requieren inteligencia humana». Se puede decir que la IA es la capacidad de las máquinas para usar algoritmos, aprender de los datos y utilizar lo aprendido en la toma de decisiones tal y como lo haría un ser humano. Sin embargo, a diferencia de las personas, los dispositivos basados en IA no necesitan descansar y pueden analizar grandes volúmenes de información a la vez. Asimismo, la proporción de errores es significativamente menor en las máquinas que realizan las mismas tareas que sus contrapartes humanas.

La IA se puede aplicar en casi todas las situaciones. Éstas son sólo algunas de las aplicaciones técnicas de la IA que están creciendo rápidamente en la actualidad:

- Reconocimiento de imágenes estáticas, clasificación y etiquetado: estas herramientas son útiles para una amplia gama de industrias.
- Mejoras del desempeño de la estrategia algorítmica comercial: ya ha sido implementada de diversas maneras en el sector financiero.
- Procesamiento eficiente y escalable de datos de pacientes: esto ayudará a que la atención

médica sea más efectiva y eficiente.

- Mantenimiento predictivo: otra herramienta ampliamente aplicable en diferentes sectores industriales.
- Detección y clasificación de objetos: puede verse en la industria de vehículos autónomos, aunque también tiene potencial para muchos otros campos.
- Distribución de contenido en las redes sociales: se trata principalmente de una herramienta de marketing utilizada en las redes sociales, pero también puede usarse para crear conciencia entre las organizaciones sin ánimo de lucro o para difundir información rápidamente como servicio público.
- Protección contra amenazas de seguridad cibernética: es una herramienta importante para los bancos y los sistemas que envían y reciben pagos en línea.

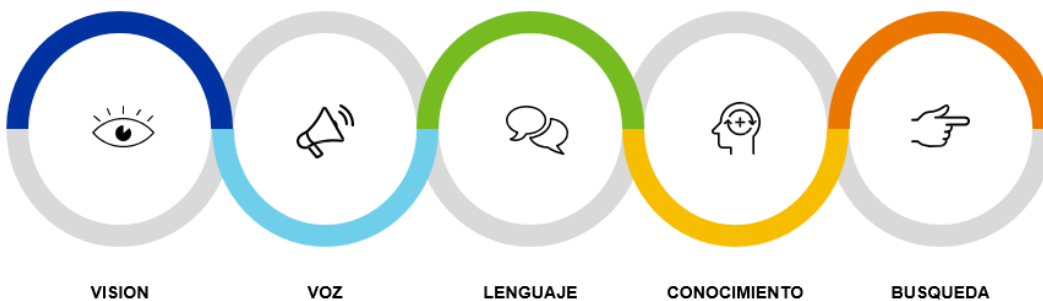


Figura 2.18: Ejemplos de áreas en donde la Inteligencia Artificial se puede usar con grandes resultados.

2.5.3. Ramas de la IA

La inteligencia artificial es una de las ramas de la informática construida fundamentalmente en áreas lógicas y de ciencias cognitivas. Debido a su naturaleza, la IA cuenta con diversos campos

de estudio, ramas con especialización que surgen a través de la innovación y el desarrollo de tecnología

Sistemas expertos

También conocidos como “Sistemas Basados en Conocimiento”, un SE es un conjunto de programas informáticos que aplican el razonamiento humano experto en la materia para la solución de problemas específicos, con estos programas es posible tomar decisiones inteligentes a partir del conocimiento y la capacidad de interpretar cantidades enormes de datos, con la función primordial de ser un auxiliar de profesionales expertos

Redes neuronales

El objetivo de las redes neuronales es tratar de emular el proceso biológico de aprendizaje que se da en los seres humanos a través de procesadores que operan de manera simultánea, se origina a través de estímulos en la red, el procesamiento local y las conexiones realizadas por medio de nodos neuronales. Se debe aclarar que estos sistemas no se pueden programar, son entrenados a través de ruidos para ser capaces de reconocer objetos.

Lógica difusa

La lógica difusa o también conocida como lógica heurística es la aplicación de otro tipo de lógica diferente a la tradicional, se basa en el cálculo de palabras y el razonamiento aproximado que trata de simplificar y resolver problemas de manera natural, eficiente y robusta. Este tipo de lógica es utilizada de manera precisa en aquellos procesos con un nivel de complejidad muy altos en los cuales no existen modelos matemáticos.

Robótica

Una parte de la robótica también está considerada dentro de las ramas de la IA, si se tiene en mente que un robot es un dispositivo electromecánico que está programado para que realice tareas manuales, estos robots utilizan el principio de la IA cuando son capaces de responder frente a estímulos sin necesidad de seguir instrucciones momentáneas, esto sucede gracias a la ayuda de cámaras y sensores que permiten interpretar las señales del entorno y actuar frente a ellos.

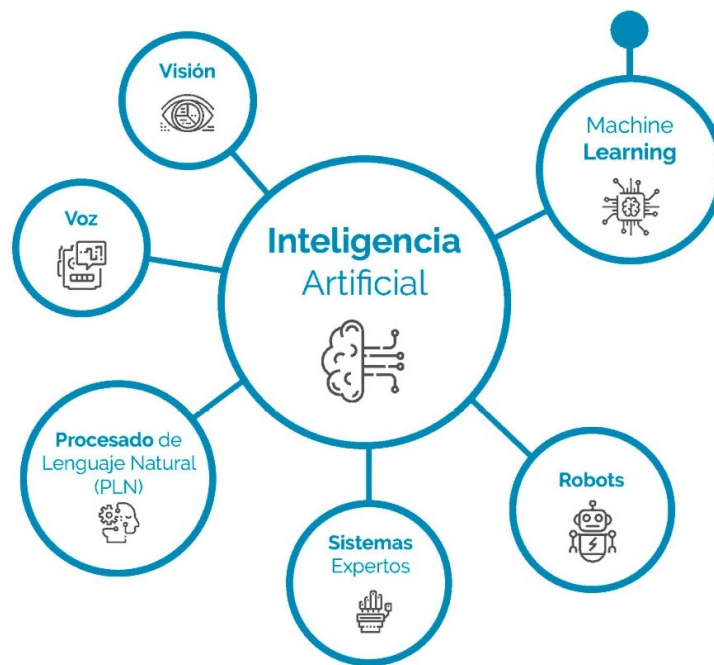


Figura 2.19: Las diferentes ramas de la Inteligencia Artificial en las cuales se puede profundizar y desarrollar grandes proyectos.

2.5.4. Aplicaciones

IA en la Publicidad

Esta herramienta permite liberar de trabajos con mucha carga de análisis procesual y computacional, para que las personas sean las encargadas de realizar trabajos que a priori son de mayor

valor comercial y que las máquinas no pueden hacer, como son la creatividad y la gestión estratégica. Decimos a priori porque la tecnología de la inteligencia artificial ya se está implementando también en las ramas de la creatividad. (Leyva-Vásquez, 2018)

Este tipo de herramienta está permitiendo desarrollar un proceso de creación publicitaria, en el que los algoritmos aportan un inmenso valor en campos como el de la programática digital, proporcionando una personalización extrema de las audiencias, análisis de datos u optimizar los resultados de las campañas y el ROI de las mismas. Estas son tareas a las que los humanos necesitan dedicar un tiempo mucho mayor que las máquinas a la hora de tomar la decisión que mejor se adecua a determinada campaña y las inteligencias artificiales son capaces de realizar todo este proceso en tiempo real y en cuestión de milisegundos.

Hay muchas esperanzas depositadas en la inteligencia artificial en cuanto a los beneficios que aporta al proceso de compraventa de publicidad e incluso la influencia que ya ejerce en el proceso de la creatividad publicitaria, es decir, la inteligencia artificial está traspasando todos los límites imaginables hace sólo algunos años, está transformando la mayoría de los ámbitos de la industria publicitaria. Esto se ve claramente reflejado en la inversión publicitaria de los Estados Unidos, donde se prevé que se destinen hasta 172.3 billones de dólares a la publicidad digital en el año 2021.

IA en Medicina

Recientemente la IA ha comenzado a incorporarse a la medicina para mejorar la atención al paciente al acelerar los procesos y lograr una mayor precisión diagnóstica, abriendo el camino para brindar una mejor atención médica en general. Las imágenes radiológicas, las preparaciones de anatomía patológica y los registros médicos electrónicos de los pacientes se están evaluando me-

diante aprendizaje automático ayudando en el proceso de diagnóstico y tratamiento de los pacientes

De esta manera existen proyectos en la actualidad dedicados a explorar las aplicaciones de la IA en todas las facetas sanitarias: asistencial (prevención de enfermedades, diagnóstico, tratamiento y seguimiento de pacientes), docente o formación continuada, investigadora y gestora. A continuación se comentan algunos ejemplos concretos en las diferentes áreas de aplicación sanitaria. (Exposito - Ávila ,2018)

1. Prevención de enfermedades y diagnóstico precoz: Existen algoritmos informáticos que son capaces de contribuir a la prevención del cáncer de cérvix con alta precisión, ya sea a través de aplicación de software de machine learning en la identificación del virus del papiloma humano o de células con transformaciones oncogénicas.

Sobre el contenido de mensajes publicados en redes sociales también se pueden generar algoritmos que predicen el riesgo de psicopatologías. Se están desarrollando estudios sobre contenido de publicaciones y su relación con procesos psicológicos como el riesgo de depresión según los contenidos de las redes sociales como Twitter o Instagram. A través de la interpretación de datos que registran nuestros movimientos (a través de sensores en nuestro teléfono móvil, reloj inteligente u otros dispositivos) se pueden identificar patrones que sugieran progresión de la enfermedad de Parkinson.

2. Diagnóstico: Existen muchos programas informáticos de apoyo y ayuda al diagnóstico que han ido mejorando su aprendizaje a través de su uso repetido y continuado. Actualmente existen diferentes tipos de software que se pueden aplicar a diferentes grupos de enfermedades como MYCIN/MYCIN II para enfermedades infecciosas, CASNET para oftalmología, PIP para enfermedades renales o AI/RHEUM para enfermedades reumatológicas.



Figura 2.20: La vision artificial siendo usada en procesos médicos para mejores diagnósticos y posteriores resultados.

En el campo del procesamiento y la interpretación de imágenes para el diagnóstico, la IA ofrece algoritmos que mejoran la calidad y la precisión del diagnóstico ya que los métodos de IA son excelentes para reconocer automáticamente patrones complejos en los datos de imágenes, elimina ruido en las imágenes ofreciendo una mayor calidad y permite establecer modelos tridimensionales a partir de imágenes de pacientes concretos

3.Tratamiento: Combinando diferentes aplicaciones tecnológicas como localización GPS, IA, sensores corporales en tejidos inteligentes o complementos de vestido podemos predecir comportamientos o actividades de personas mayores que viven solas pudiendo mejorar su autonomía.

La IA también se puede aplicar para predecir reacciones adversas de tratamientos médicos o el grado de cumplimentación del tratamiento por parte de los pacientes.

Capítulo 3

METODOLOGÍA

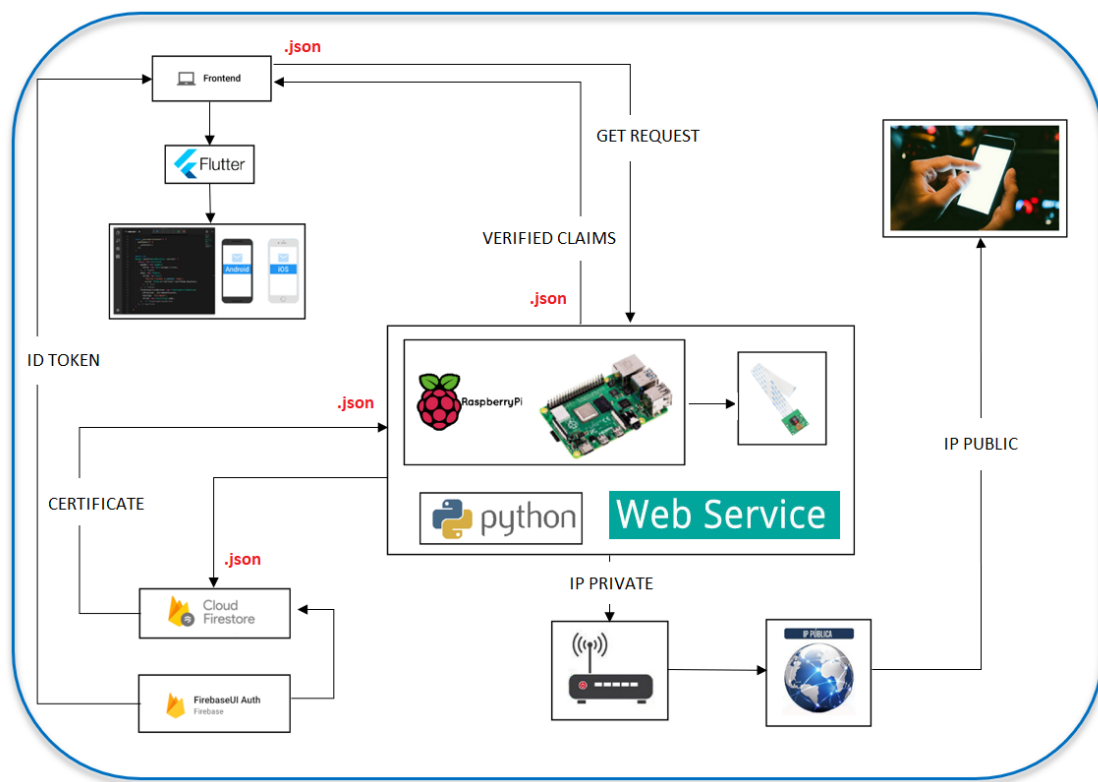


Figura 3.1: Arquitectura general del proyecto evidenciando claramente las 4 subdivisiones importantes (Backend, Frontend, Reconocimiento facial y App final).

3.1. Normativas, leyes, reglamentos

El estado colombiano en su constitución presenta un conjunto de normativas y leyes respecto a la manipulación de datos y uso de información personal que deben ser tenidos en cuenta cuando se realizan proyectos de aprendizaje automático aplicados a personas que, naturalmente, deben trabajar con datos de diferentes personales y su uso desautorizado puede terminar en problemas legales.

Ley 1581 de 2012

La Ley 1581 de 2012 estipula un marco legal en el tratamiento de datos personales en Colombia, estableciendo los usos que se puede dar a la información privada de una persona bajo su consentimiento, como dicha persona puede rectificar la información que se da en una base de datos y como se debe castigar a quienes usan la información personal sin autorización. En este sentido, la ley enuncia que:

Recolección de los datos personales. En desarrollo de los principios de finalidad y libertad, la recolección de datos deberá limitarse a aquellos datos personales que son pertinentes y adecuados para la finalidad para la cual son recolectados o requeridos conforme a la normatividad vigente. Salvo en los casos expresamente previstos en la Ley, no se podrán recolectar datos personales sin autorización del Titular.

El responsable del Tratamiento deberá adoptar procedimientos para solicitar, a más tardar en el momento de la recolección de sus datos, la autorización del Titular para el Tratamiento de los mismos e informarle los datos personales que serán recolectados así como todas las finalidades

específicas del Tratamiento para las cuales se obtiene el consentimiento.

Los datos personales que se encuentren en fuentes de acceso público, con independencia del medio por el cual se tenga acceso, entendiéndose por tales aquellos datos o bases de datos que se encuentren a disposición del público, pueden ser tratados por cualquier persona siempre y cuando, por su naturaleza, sean datos públicos.

3.2. Backend

El back end de la app consiste en un servidor y una base de datos. Se toman los datos, se procesa la información y se envía al usuario. Los desarrolladores de Front end y Back end suelen trabajar juntos para que todo funcione correctamente. Un desarrollador Back end debe tener amplios conocimientos de los siguientes lenguajes: frameworks y los tipos de base de datos. No siendo necesario conocer todos los lenguajes pero sí entender y saber trabajar con algunos de ellos.

Para la configuración del servidor se trabajó de dos maneras: En la primera, se configuró el algoritmo de inteligencia artificial en una tarjeta de desarrollo llamada Raspberry Pi con su módulo de cámara; y la segunda un alojamiento de la app móvil en un servidor Google. Para entender más a fondo primero se aborda unos pequeños conceptos a tener en cuenta como la definición de la raspberry, su módulo de cámara y la base de datos Firebase.

3.2.1. Raspberry Pi 4 B

Raspberry Pi es, básicamente, un ordenador. A pesar de su placa reducida, su bajo coste y su sistema operativo de código abierto Raspbian, no deja de ser un ordenador adaptado a unas ne-

cesidades de programación muy básicas. Diseñados en un principio para fomentar la enseñanza informática en las escuelas, ha tenido mucho éxito durante sus sucesivas evoluciones.

El motivo de su gran popularidad no tiene tanto que ver con su utilidad académica (que también), sino por las capacidades emergentes que se descubren continuamente de él. Para ser un ordenador de aproximadamente 550.000 COP en su ultimísima versión, no es nada desdeñable poder utilizarlo de centro multimedia, como controlador o incluso como servidor de juego en Minecraft.

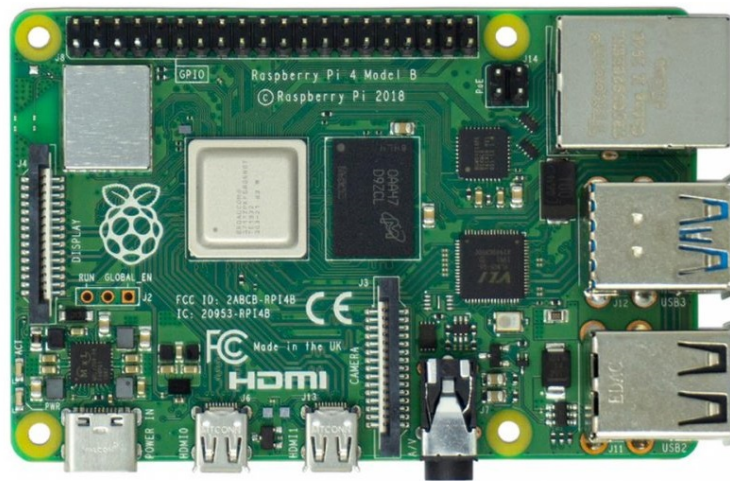


Figura 3.2: Imagen de una Raspberry Pi 4B 8 GB desde la capa superior donde se pueden ver algunos de sus componentes.

Su uso más obvio, y quizá por el que fue creado en un primer momento, es el de ser un ordenador de escritorio que cumpla las funciones más básicas. En este caso, se necesitan los periféricos (teclado y ratón), así como un monitor al que enchufar un cable HDMI, una tarjeta SD y la versión del sistema operativo que queramos. A partir de ahí, podremos hacer las tareas diarias más simples como navegar por internet, revisar correos o redactar textos en un PC que apenas ocupa espacio.



Figura 3.3: Raspberry conectado con periféricos siendo usada como ordenador de bajo coste y aceptables prestaciones.

En el caso de incorporar una cámara fotográfica a nuestra Raspberry Pi 4, nos topamos con un abanico enorme de posibilidades que ha ido desarrollando la comunidad. Y es que un aspecto tan artístico como la fotografía y todas las funcionalidades que ella conlleva, tiene un potencial enorme si se cuenta con un recurso tan útil como este mini ordenador. Así, ya sea mediante Wi-Fi u otras vías posibles, podemos configurar la secuencia de capturas de nuestras cámaras para automatizar los procesos propios de recursos tales como Time Lapse o Slow Motion.

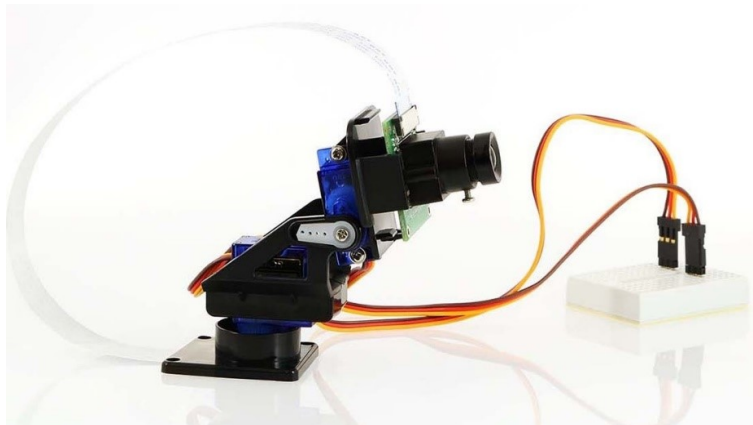


Figura 3.4: Raspberry usada en visión artificial conectada a servomotores para aplicaciones de reconocimiento de objetos o facial.

En definitiva, Raspberry Pi 4 cuenta con tantos usos que puede llegar a satisfacer hasta al menos

interesado en informática. Ya sea por precio, versatilidad o la simple curiosidad que despierta, este ordenador supone un elemento que refuerza las posibilidades de crear de cualquier persona con inquietudes. Y pensar que esto era para las escuelas.

En la figura 3.5 se muestra las características de 3 modelos diferentes de Raspberry Pi. Para este trabajo se hizo uso de la Modelo 4 B de 8 GB de memoria RAM.

	RASPERRY PI 4 MODEL B	RASPERRY PI 3 MODEL B+	RASPERRY PI 3 MODEL B
PROCESADOR	Broadcom BCM2711B0, quad-core Cortex-A72	Broadcom BCM2837B0, Cortex-A53 (ARMv8) 64-bit SoC	Broadcom BCM2837, Cortex-A53 (ARMv8) 64-bit SoC
FRECUENCIA DE RELOJ	1,5 GHz	1,4 GHz	1,2 GHz
GPU	VideoCore VI 500 MHz	VideoCore IV 400 MHz	VideoCore IV 400 MHz
MEMORIA	1/2/4/8 GB LPDDR4-3200	1GB LPDDR2 SDRAM	1GB LPDDR2 SDRAM
CONECTIVIDAD INALÁMBRICA	Wi-Fi 2,4GHz / 5GHz IEEE 802.11.b/g/n/ac Bluetooth 5.0, BLE	Wi-Fi 2,4GHz / 5GHz IEEE 802.11.b/g/n/ac Bluetooth 4.2, BLE	Wi-Fi 2,4GHz IEEE 802.11.b/g/n Bluetooth 4.1
CONECTIVIDAD DE RED	Gigabit Ethernet	Gigabit Ethernet over USB 2.0 (300 Mbps de máximo teórico)	Fast Ethernet 10/100 Gbps
PUERTOS	GPIO 40 pines 2 x Micro HDMI 2 x USB 2.0 2 x USB 3.0 CSI (cámara Raspberry Pi) DSI (pantalla tácil) Toma auriculares / vídeo compuesto Micro SD USB-C (alimentación) Power-over-Ethernet (PoE)	GPIO 40 pines HDMI 4 x USB 2.0 CSI (cámara Raspberry Pi) DSI (pantalla tácil) Toma auriculares / vídeo compuesto Micro SD Micro USB (alimentación) Power-over-Ethernet (PoE)	GPIO 40 pines HDMI 4 x USB 2.0 CSI (cámara Raspberry Pi) DSI (pantalla tácil) Toma auriculares / vídeo compuesto Micro SD Micro USB (alimentación)
FECHA DE LANZAMIENTO	24/06/2019	14/3/2018	29/2/2016
PRECIO	39,87 / 50,76 / 60,44 euros	39,75 euros	33,95 euros

Figura 3.5: Comparación de varias características y parametros importantes entre modelos de Raspberry Pi.

Al entender de mejor manera que es una tarjeta o minicomputador como lo es la Raspberry Pi, se puede mencionar que se hizo dentro de ella. Este pequeño PC cuenta con un desarrollador Python en el cual se programó los dos metodos de clasificador de rostros, tanto por inteligencia

artificial como por OpenCV. En la figura 3.7 se evidencia la interfaz grafica de la Raspberry Pi dentro de un PC anfitrión, ya que se hizo uso de conexión SSH para trabajar desde otro PC y no con monitor y periféricos externos.

Este miniordenador se adquirió nuevo por lo que primero se debe realizar la instalación de su SO descargado desde su sitio web oficial <https://www.raspberrypi.org/> , luego de instalarlo correctamente se realiza una configuración normal de red local.

Esta configuración se realiza con el ámbito de poder conectarla a la red e ingresar desde otro ordenador tanto para configurarla como para la administración de su módulo de cámara. Para visualizar el escritorio se hace uso del programa VNC Viewer activando previamente vncserver en la pantalla de comandos por Putty.

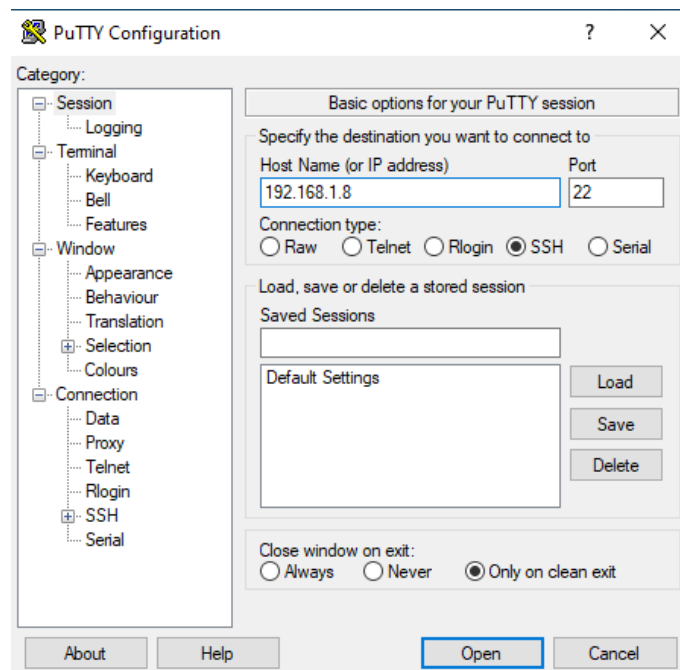


Figura 3.6: Conexión por el protocolo SSH entre la Raspberry Pi usada y el PC mediante el software Putty.

Luego de tener la conexión activa y funcionando perfectamente la Raspberry Pi Camera Board

se conecta directamente al conector CSI de la Raspberry Pi. Esta cámara es capaz de ofrecer una imagen nítida de 5MP de resolución, o una grabación de vídeo HD de 1080p a 30fps. Diseñada y fabricada por la Fundación Raspberry Pi en el Reino Unido, la Raspberry Pi Camera Board cuenta con un sensor Omnivision 5647 de 5MP (2592x1944 píxeles) en un módulo de enfoque fijo. El módulo se conecta a la Raspberry Pi, por medio de un cable de cinta de 15 pines, a la interfaz en serie de la cámara MIPI de 15 pines (CSI), que fue diseñada especialmente para la interfaz con las cámaras. El bus CSI es capaz de alcanzar velocidades de datos extremadamente altas, y transporta exclusivamente los datos de los píxeles al procesador BCM2835.

La placa en sí es diminuta, de unos 25 mm x 20 mm x 9 mm, y pesa poco más de 3 g, lo que la hace perfecta para aplicaciones móviles o de otro tipo en las que el tamaño y el peso son importantes. El sensor tiene una resolución nativa de 5 megapíxeles y cuenta con un objetivo de enfoque fijo. En cuanto a las imágenes fijas, la cámara es capaz de tomar imágenes estáticas de 2592 x 1944 píxeles, y también admite la grabación de vídeo a 1080p @ 30fps, 720p @ 60fps y 640x480p 60/90. La cámara es compatible con la última versión de Raspbian, el sistema operativo preferido de Raspberry Pi.

Características:

- Totalmente compatible con el modelo A y el modelo B de Raspberry Pi.
- Módulo de cámara Omnivision 5647 de 5MP.
- Resolución de imagen fija: 2592 x 1944.
- Vídeo: Soporta 1080p @ 30fps, 720p @ 60fps y 640x480p 60/90 Grabación.

- Interfaz en serie de la cámara MIPI de 15 pines: se conecta directamente a la placa Raspberry Pi.
- Tamaño: 20 x 25 x 9mm.
- Peso 3g
- Totalmente compatible con muchas carcasas de Raspberry Pi

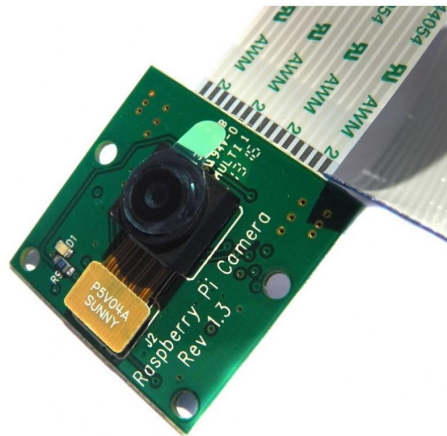


Figura 3.7: Vista lateral de módulo de Raspberry Pi Camera Board v1.3. Módulo utilizado en este proyecto.

Al comprobar la conexión y buen funcionamiento de la cámara tanto para la toma de fotografías como para la toma de videos en una buena calidad, se procede al desarrollo de la app móvil.

La mayoría de las apps necesitan identificar a los usuarios. Conocer la identidad de un usuario permite que una app guarde sus datos en la nube de forma segura y proporcione la misma experiencia personalizada en todos los dispositivos del usuario. Firebase Authentication proporciona servicios de backend, SDK fáciles de usar y bibliotecas de IU ya elaboradas para autenticar a los usuarios en tu app. Admite la autenticación mediante contraseñas, números de teléfono, proveedores de identidad federada populares, como Google, Facebook y Twitter, y mucho más.

Firebase Authentication se integra estrechamente con otros servicios de Firebase y aprovecha

estándares de la industria como OAuth 2.0 y OpenID Connect, por lo que se puede integrar con facilidad en tu backend personalizado.



Figura 3.8: Logo Firebase Authentication. Backend utilizado para el login de los usuarios de la app creada.

¿Cómo funciona?

Para que un usuario acceda a la app, primero se debe obtener credenciales de autenticación del usuario. Estas credenciales pueden ser la dirección de correo electrónico y la contraseña del usuario, o un token OAuth de un proveedor de identidad federada. Luego, se debe pasar estas credenciales al SDK de Firebase Authentication. Después, los servicios de backend verificarán esas credenciales y mostrarán una respuesta al cliente.

Cuando el acceso se realiza correctamente, se puede ver la información básica del perfil del usuario y controlar el acceso del usuario a los datos almacenados en otros productos de Firebase. También se puede utilizar el token de autenticación proporcionado para verificar la identidad de los usuarios en tus propios servicios de backend.



Figura 3.9: Modos de entrada que permite Firebase Auth para ingresar a la aplicación (Facebook, Google, Twitter, entre otras.)

Funciones clave:

- Se integra con otras funciones de Firebase.
- Usa estándares de la industria, como OAuth 2.0 y OpenID Connect, para que pueda integrarse fácilmente a tu sistema back-end personalizado.
- Ofrece dos opciones de desarrollo, ya sea FirebaseUI como una solución de autenticación fácil de implementar o el SDK de Firebase Authentication para integrar de forma manual uno o varios métodos de acceso a tu app.
- Brinda autenticación segura que facilita el acceso de los usuarios mediante su Cuenta de Google con la que ya usan Gmail, Google Play y otros servicios de Google. También admite la autenticación mediante contraseñas y proveedores de identidad populares, como Facebook y Twitter.
- Ofrece una experiencia en la app sin dificultades entre dispositivos y en tu sitio web de manera segura y con un consentimiento único. De esta forma, mantendrás comprometidos a tus usuarios independientemente del dispositivo que utilicen.

- Conecta a los usuarios de forma segura con los servicios de Google. Comparte con Contactos de Google, guarda archivos en Drive, agrega eventos a Calendario y mucho más.
- Permite a los usuarios pagar con Google Wallet. Ofrece la posibilidad de pagar con Google Wallet, lo que simplifica las compras directas desde la aplicación.

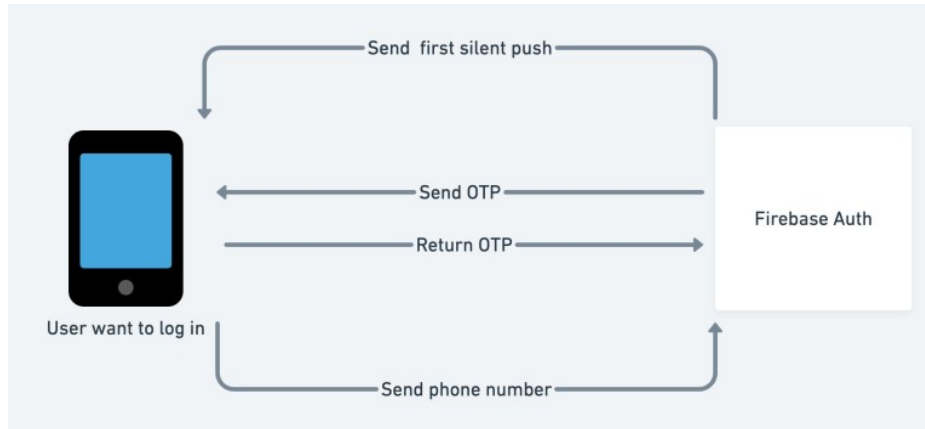


Figura 3.10: Esquema del funcionamiento de Firebase Authentication entre el servidor y el móvil.

3.3. Frontend

3.3.1. Diseño de la aplicación móvil

Al tener definidos todos los parámetros y programación que tendrá el proyecto se debe pensar en cómo los usuarios finales tendrán acceso a estos parámetros. Una forma eficaz y que está en vanguardia en la actualidad son las aplicaciones móviles. Es muy común que todas las personas actualmente dispongan de un teléfono celular tanto para fines recreativos como para laborales, y en dichos teléfonos vemos aplicaciones de todo tipo. Un problema notorio es que todos los usuarios no tienen la misma capacidad en conocimiento de manejo de aplicaciones móviles, por lo tanto, se hace necesario el diseño de una aplicación sencilla y de fácil uso para la interacción entre el

backend y el usuario final.

Otro item muy importante a tener en cuenta es que actualmente los teléfonos celulares no tienen un único sistema operativo. Los sistemas operativos usados para los teléfonos móviles, celulares o smartphone son muchos, pero hay 2 que son los principales y que ocupan casi todo el mercado de la telefonía móvil: El Android y el iOS. Seguidos, pero con mucha diferencia, estarían Symbian, Blackberry OS y Windows Phone. Para ello se hace necesario que el desarrollo de la aplicación móvil sea en una plataforma capaz de soportar los dos sistemas operativos, en este caso se escoge Visual Code con un framework llamado FLUTTER.

Flutter es el conjunto de herramientas de interfaz de usuario de Google para crear hermosas aplicaciones compiladas de forma nativa para dispositivos móviles, web y de escritorio a partir de una única base de código.

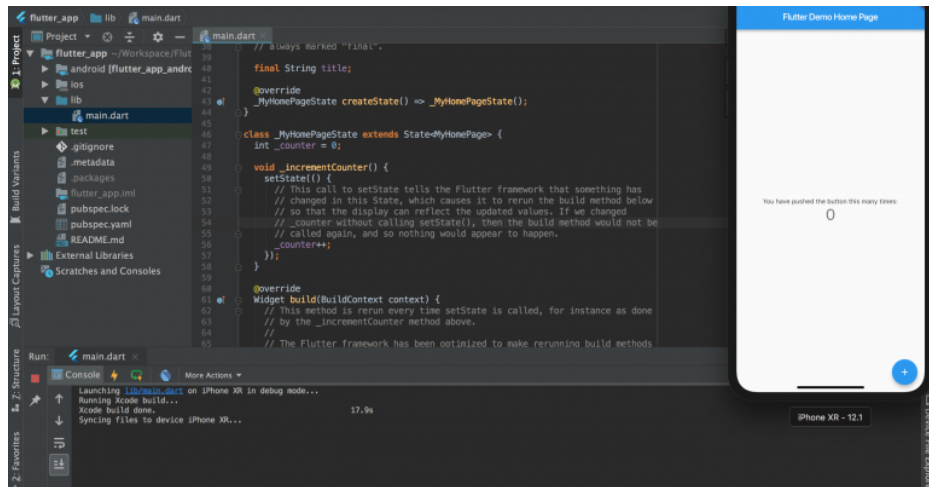


Figura 3.11: Interfaz gráfica de Visual Studio Code, donde se trabaja el framework Flutter y su lenguaje nativo es Dart.

En primera instancia, se crea una pantalla a la que llamaremos “Splash”. Esta pantalla estará constituida por 3 widgets. Abordemos primero el concepto de Widget. Una definición rápida de

internet: “En informática, un widget o artilugio es una pequeña aplicación o programa, usualmente presentado en archivos o ficheros pequeños que son ejecutados por un motor de widgets o Widget Engine. Entre sus objetivos están dar fácil acceso a funciones frecuentemente usadas y proveer de información visual.”

Ya entendiendo que es un widget, haremos uso de tres de ellos, en este caso serán 3 imágenes contenidas en un `SizedBox`, el cual es una especie de contenedor o caja donde se puede colocar contenido visible en la aplicación. Estas 3 imágenes serán: el escudo de la Universidad de Pamplona, el escudo del programa de Ingeniería en Telecomunicaciones de la Universidad de Pamplona, y el logo de la aplicación a la cual llamaremos JESI.

El nombre se define por sus siglas en ingles: “Just a Environment Safe and Intelligent” el cual traduce: “Solo un entorno seguro e inteligente”. Esto debido a que la programación tanto en dart como en Python se desarrollaron en entornos que se crearon para la no influencia de las variables y paquetes en el host donde se creó dicha aplicación.

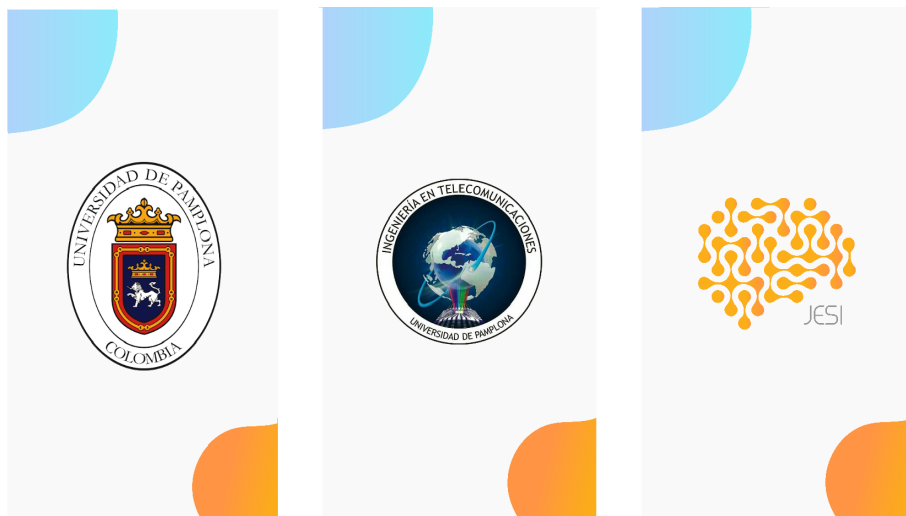


Figura 3.12: Pantallas con los escudos de la Universidad de Pamplona, del programa Ingeniería en Telecomunicaciones y del logo de la app para crear un splash de inicio.

En la figura 3.12 se puede apreciar las 3 pantallas que se usarán en nuestra pantalla Splash. Al tenerlas ya listas lo que procede es hacer una transición automática entre las 3 pantallas con efecto de desvanecido de las imágenes para una mejor apariencia. Esta pantalla Splash aparecerá en el momento en que la aplicación inicie, en términos sencillos, se podría hablar de la pantalla de inicio o presentación de la aplicación móvil.

Luego de esta pantalla de presentación, aparece una de las pantallas o screens mas importantes de la aplicación, la pantalla de autenticación o logueo. Esta pantalla esta dividida en dos partes, en la primera se encuentra de nuevo el logo de la aplicación como método de presentación; y en la segunda, la mas importante, se encuentra los campos de autenticación.



Figura 3.13: Pantalla de logueo de la app con sus respectivos campos de usuario y contraseña.

En esta segunda parte, se hizo uso de 2 widgets llamados `TextInputField`, los cuales son un campo que reciben texto, parte fundamental en el logueo. Luego se tiene un botón para iniciar sesión que llevará al “Home” de la aplicación si los datos son correctamente validados.

Debajo de estos dos `TextPutField` encontramos un texto el cual nos sirve para que los nuevos usuarios se registren. Este registro se desarrolló de una manera fácil para el usuario, solo debe colocar su correo electrónico, una contraseña mayor a 6 caracteres y dar en “AQUÍ”. De esta manera, la aplicación se conectara con `Firebase Auth` y guardara su usuario y contraseña para futuros inicios de sesión. En la figura 3.14 se muestra esta pantalla de logueo, cabe recordar que todas estas pantallas han tenido un fondo predeterminado creados con anterioridad para una buena apariencia jugando con la psicología del color.

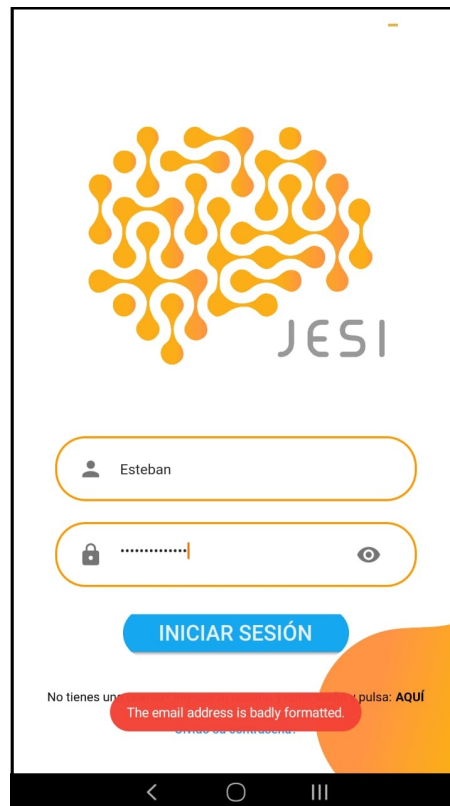


Figura 3.14: Mensaje de alerta cuando el usuario ingresado no corresponde a alguno que se encuentra en la base de datos.

En el primer TextField se ingresa el usuario, cabe resaltar que en el algoritmo desarrollado se especifica que el usuario debe ser tipo e-mail, de no ser así, se presentará un mensaje tipo burbuja mostrando este error. Esto se evidencia en la figura 3.15.

En el segundo TextField se ingresa la contraseña de tipo numérica, alfabética o alfanumerica, la unica condición que se presenta es que tenga mas de 6 caracteres,de no ser asi, se entregará otro mensaje burbuja indicando que la contraseña es invalida. El último mensaje burbuja que puede aparecer es cuando la contraseña no corresponde al usuario escrito en el campo usuario”, para ello se debe rectificar e ingresar nuevamente la contraseña.En la figura 3.16 se puede observar estos dos mensajes burbujas.

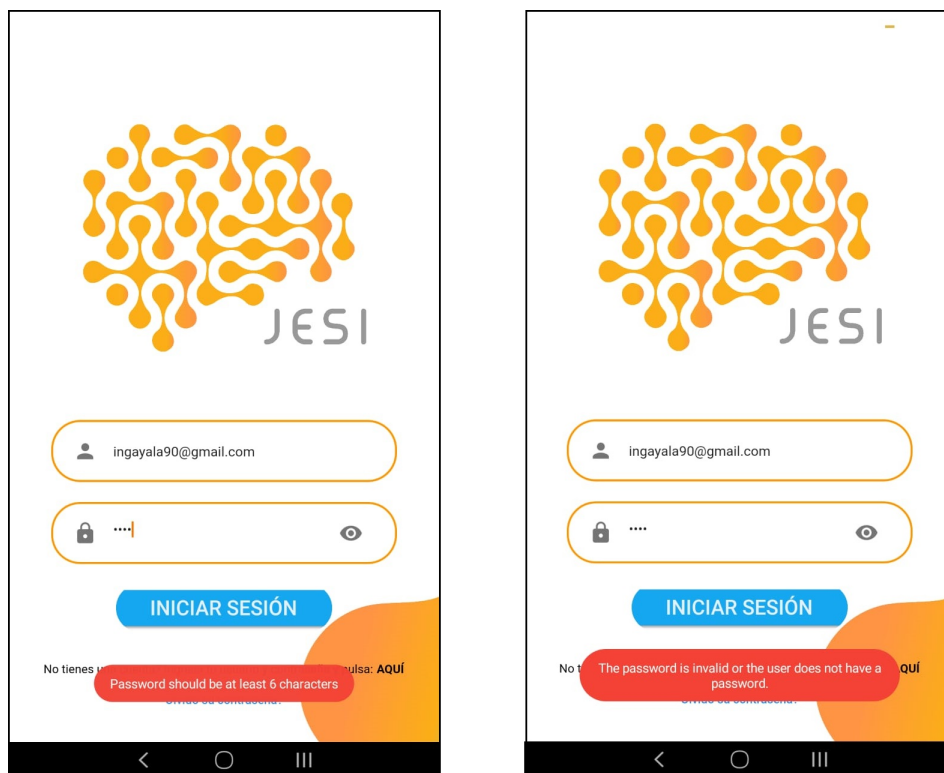


Figura 3.15: A la izquierda, Mensaje de error cuando la contraseña contiene menos de 6 caracteres. A la derecha, error de Contraseña invalida para el usuario.

Una opción que no se debe olvidar y hay que tener en cuenta es cuando al usuario por una u otra razón olvida su contraseña y necesita recobrarla. Para ello, debajo del texto de registro se ubica un boton con la leyenda OLVIDÓ SU CONTRASEÑA, al entrar ahí nos llevará a una nueva ventana donde debemos ingresar nuestro usuario y luego dar click en el boton RESET. Esto se puede evidenciar de forma clara en la figura 3.18.

Al dar click en RESET, en un lapso de no mas de 3 min. llegará un correo electrónico, con un vinculo que guiará para establecer una nueva contraseña.(Ver figura 3.17)

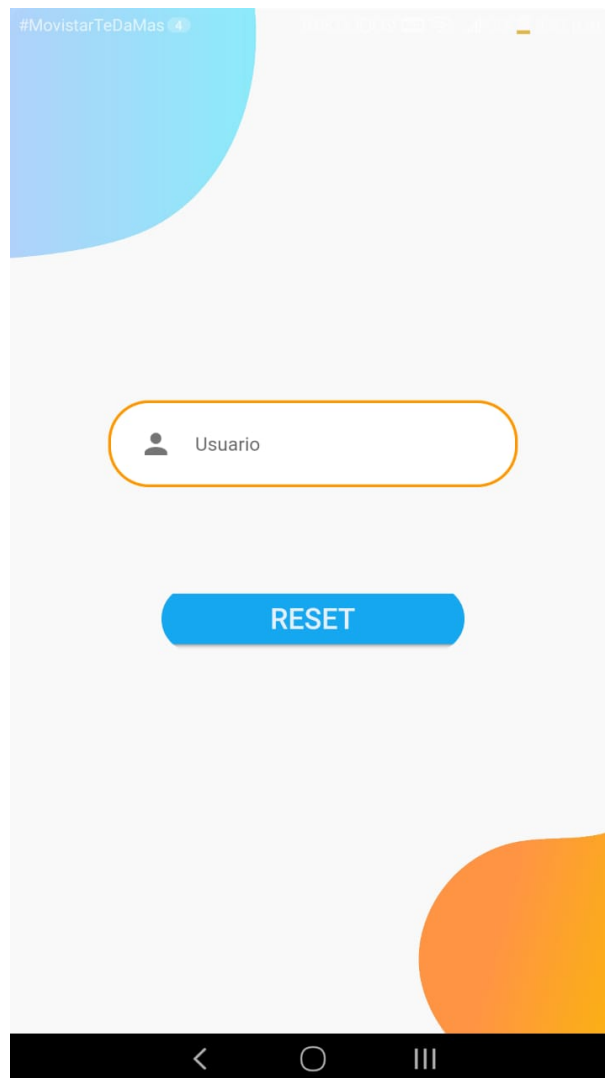


Figura 3.16: Pantalla para ingreso de usuario el cual quisiera hacer reset a su contraseña olvidada.

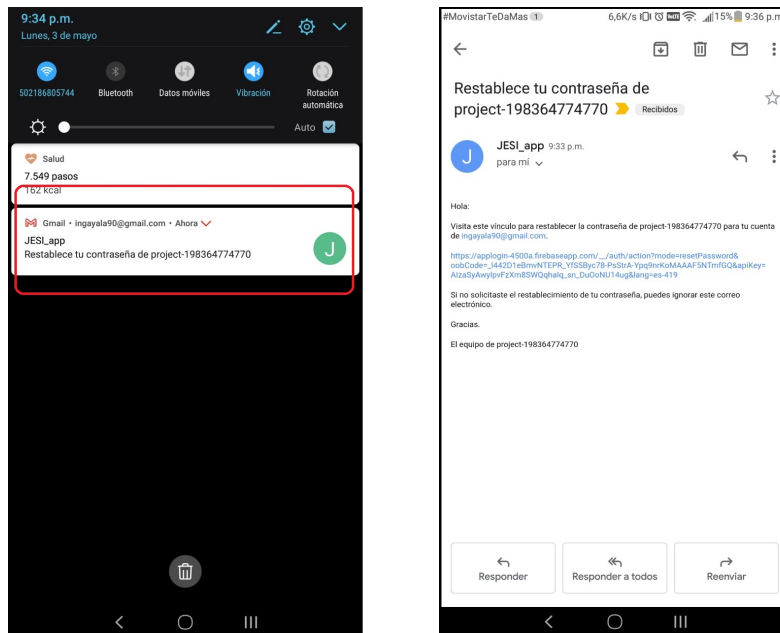


Figura 3.17: A la izquierda, Notificación de Llegada de e-mail de la app para reestablecer contraseña. A la derecha, contenido del correo para reestablecer dicha contraseña.

Al momento de dar click en el vinculo se abrirá una pagina web donde se podrá colocar la nueva contraseña. Esta pagina web es de un entorno sencillo y fácil de dominar. Luego de ingresar la nueva contraseña se debe oprimir el botón GUARDAR y así se actualizarán los datos del usuario.



Figura 3.18: Pantalla luego de dar clic al correo con su campo para introducir la nueva contraseña y poder guardarla.

Si el correo de recobro de la contraseña se deja inactivo mas de una hora o el enlace ya fue usado, este mostrará un mensaje de error lo que indica que de nuevo se debe solicitar un vinculo de resetear la contraseña.(Figura 3.19)

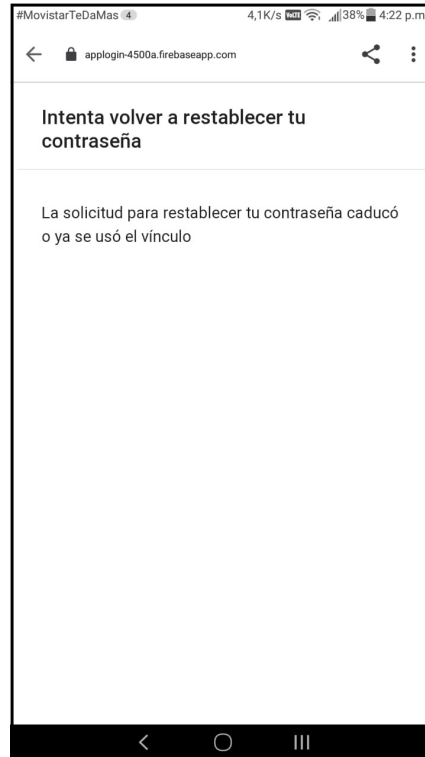


Figura 3.19: Vinculo caducado al querer recobrar la contraseña ya sea por uso anterior o por tiempo agotado para usarlo.

Ahora bien, si los datos son validados se ingresa a la pantalla principal de la aplicacion, el HOME. Este HOME se puede definir como la pantalla o screen donde se tendran todas las opciones que maneja la aplicacion adjunto de un mensaje de bienvenida.

En la parte inferior se encuentra una barra de navegación con cuatro (4) iconos. Por medio de esta barra nos podremos mover en todo el entorno de la aplicación ingresando a los 3 complementos principales: El conversor audio a texto, la camara streaming y el registro de las nuevas personas.

Una particularidad es que dependiendo de la pantalla donde se ubique el usuario el icono de la barra

inferior se iluminará de un color oscuro dejando los otros iconos en blanco, así se tendrá una mejor ubicación dentro de la arquitectura de la app.

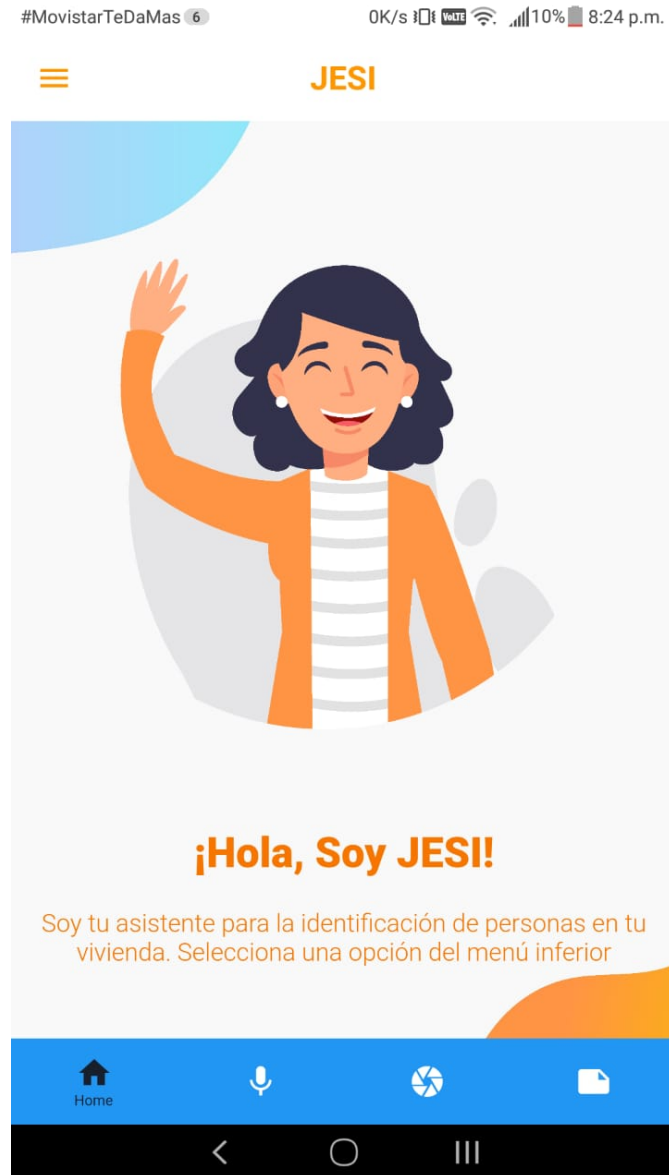


Figura 3.20: Pantalla principal de la app: el home. En ella se encuentra el menú para ingresar a todas las opciones que permiten interactuar de manera correcta.

En la parte superior o appbar se encontrará el nombre de la pantalla en donde se ubique el usuario, en este primer caso aparecerá JESI. En el appbar también se encuentra un drawer o menú. Este menú se despliega al dar click en las 3 barras horizontales donde se podrá tanto ingresar a las

notificaciones que nos genere la app y también la opción de salir y volver a la pantalla de logueo.

Esto se evidencia en la figura 3.22

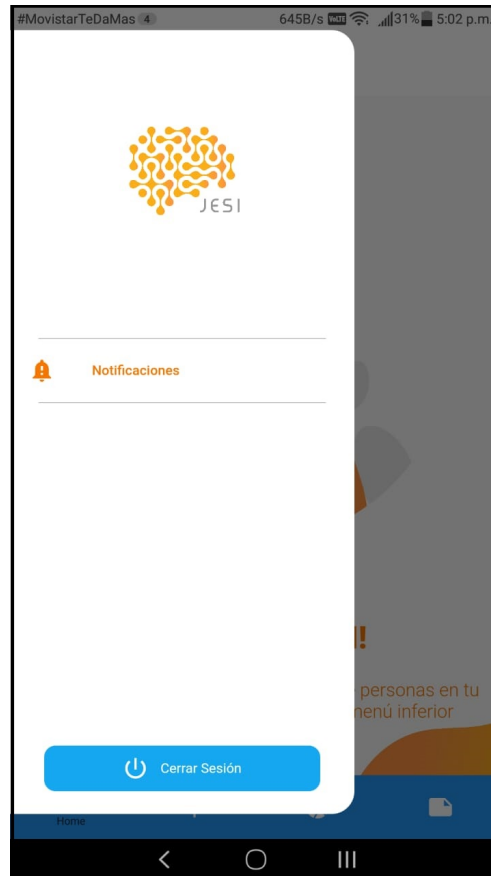


Figura 3.21: Menú lateral o drawer donde se observa el panel de notificaciones y el botón para cerrar la sesión y salir.

Entrando en la primera opción del menú inferior se navega al Screen VOZ. La función principal de este Screen o pantalla es recibir las ordenes de acciones a realizar mediante comandos de voz, pero para que esto sea posible, se tiene que analizar más a fondo como trabaja este método.

En el caso particular de esta aplicación, se hizo uso de una librería especial de Google para flutter llamada Speech to Text. Esta librería combinada con otros fundamentos logra que la aplicación pueda escuchar el entorno y realiza una comparación con su base de datos de las palabras para otorgar una respuesta lo más efectiva posible.

Ya al tener claro como funciona este proceso, se puede realizar una descripción de la pantalla. Tiene un container o cuadro de texto en la parte superior donde se mostrará la orden a realizar. Debajo de dicho container se ubica una animación que cambia dependiendo del momento en que este la app. Cuales momentos?. La pantalla tiene dos momentos especificos, el primero es cuando se activa para escuchar el entorno, y el segundo es cuando se quiere parar la grabacion de la voz y por ende, no escuche nada del entorno. Para diferenciar estos momentos, la animacion tendrá un mensaje **TE ESTOY ESCUCHANDO** para el primer momento y un mensaje de **NO TE ESCUCHO** para el segundo. Esto se explica mas claramente en las figuras 3.22 y 3.23. El control de los momentos lo tiene un boton en la parte inferior con la palabra **GRABAR**, no es necesario tener oprimido el botón, si no que con un click en el se pasa de un momento a otro.

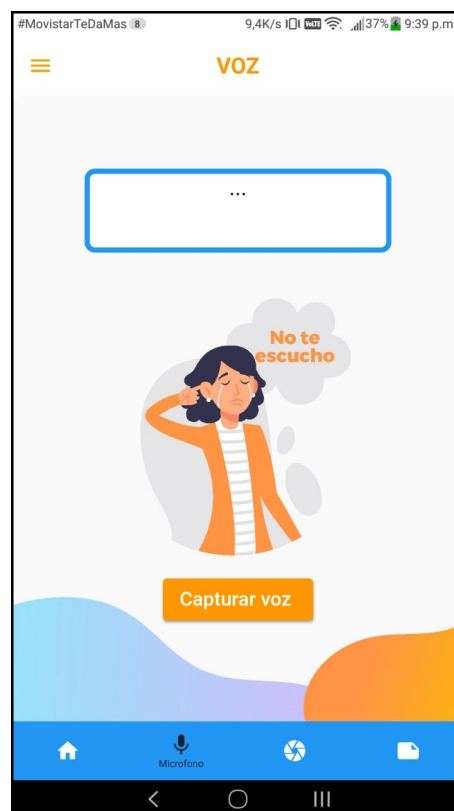


Figura 3.22: Primer momento en la pantalla de voz cuando no esta activo el listening por lo tanto no va a tomar ninguna orden.

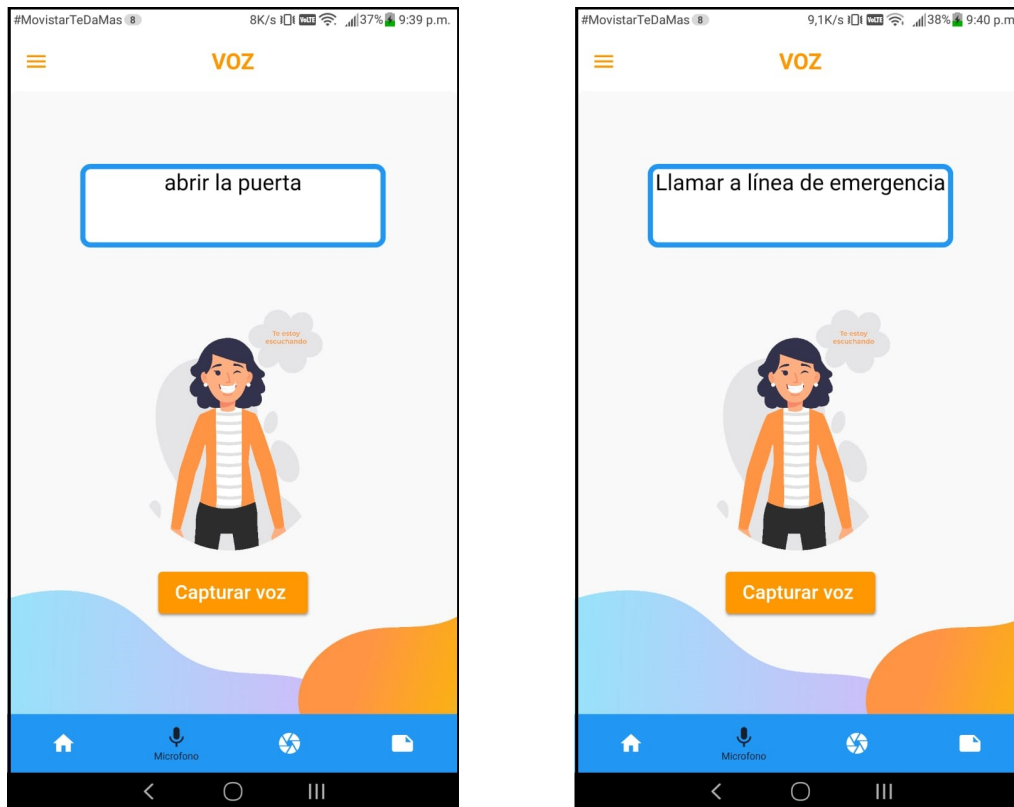


Figura 3.23: Segundo momento en la pantalla de voz cuando se puede dar ordenes por voz y al app traducirá esta voz a texto.

Como segunda opción en el menú inferior se encuentra la pantalla de la cámara Stream. Esta pantalla se puede catalogar como la mas importante de todas ya que comunica la aplicacion móvil con la camara de la Raspberry Pi ubicada en la mejor localizacion que considere el usuario administrador.

Esta pantalla cuenta en su cuerpo principal con un recuadro donde se podrá observar el contenido de la camara en tiempo real. Esta conexion streaming se realiza por medio de la libreria Flask en Python y se accede por una URL. En este caso la URL es `http://192.168.1.200`, la cual es la dirección IP de la raspberry Pi logrando asi entrar a su contenido. Además del cuadro visualizador, se encuentra en el cuerpo de la pantalla un botón con un icono en forma de X. Este boton activa y desactiva la visualización de la cámara, pero no cierra la transmisión.

Entrando a la componente mas importante de la aplicación, cuando se activa la camara y encuentra un rostro, esta realiza unos algoritmos internos encerrando en un recuadro el rostro de la o las personas que la camara encuentre identificandolas y mostrando su nombre debajo de dicho recuadro. Con esto, el usuario administrador podrá tomar las acciones pertinentes por comandos de voz, pantalla explicada anteriormente. En la figura 3.26 se puede observar lo anterior mencionado identificando cada uno de sus componentes.

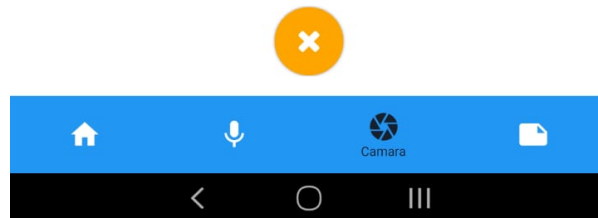
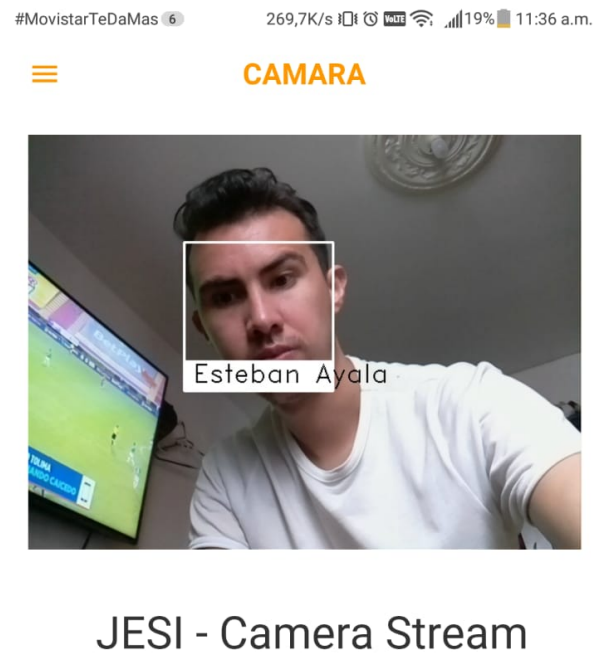


Figura 3.24: Ejemplo de detección de rostro en la app móvil con el usuario desarrollador mostrando el buen funcionamiento.

Como ultima opción en el menú inferior se encuentra el registro de las personas que podrá reconocer la aplicacion mediante la camara. Estos datos se mostrarán cuando el administrador lo solicite mediante el comando de voz: Datos de X, donde X representa el usuario que el algoritmo detecte por medio de la camara.

Los datos que se pedirán en el formulario son los nombres, al cual se le da una longitud maxima de 25 caracteres; Primer apellido: longitud maxima 12 carac.; Segundo apellido: longitud maxima 12 carac.; El numero de identificacion sea CC o TI tambien con una longitud maxima de 12 caract.; Luego una dirección de residencia, celular y RH. Hay que decir que los campos Celular y Numero de identificacion se desarrollaron de tal forma que solo acepten valores numericos, en cualquier otro caso arrojará un error para volver a ingresar dichos valores.

The image displays two screenshots of the 'REGISTRO' (Registration) screen in a mobile application. The left screenshot shows the registration form with the following fields and their maximum character counts:

- Nombres: 0/25
- Primer Apellido: 0/12
- Segundo Apellido: 0/12
- N° de identificación: 0/12
- Dirección de residencia: 0/40
- Celular: 0/10
- RH: 0/10

The right screenshot shows the same form with a 'Registrar' button and a privacy notice:

Autorizo de manera libre,previa inequivoca informada por JESI, mediante la imposición de mis datos a realizar el cotejo de mi información contra la base de datos de la Registraduría Nacional del Estado Civil (RNEC), con el fin de validar mi identidad.

Manifiesto que he sido informado que no estoy obligado a autorizar el tratamiento de dichos datos y que tengo derecho a conocer,actualizar y rectificar los datos personales proporcionados, a presentar quejas ante la Superintendencia de Industria y Comercio por el uso indebido de mis datos personales y a revocar esta autorización en los casos que sea procedente de acuerdo con la ley.

Figura 3.25: Pantalla de registro de nuevos usuarios, los cuales seran los que el usuario root no conozca para que luego la app los reconozca mediante la cámara.

Luego de llenar exitosamente todos los campos del formulario se tiene un texto de aceptación para el tratamiento de datos personales el cual el usuario administrador acepta para hacerse cargo por cualquier situación delictiva con dichos datos, posteriormente se da click en el botón inferior REGISTRAR. Estos datos quedan almacenados en una base de datos creada en Cloud Firestore. Los datos viajan en tiempo real para ser almacenados inmediatamente en su base de datos y poder ser adquiridos desde ese momento. En la figura 3.28 se observa como queda almacenados los datos de una persona en la base de datos.

Se crea un árbol el cual la raíz se denominará USUARIOS, luego debe tener un ID y dentro ese ID se ubican los campos que llegaron de la aplicación.

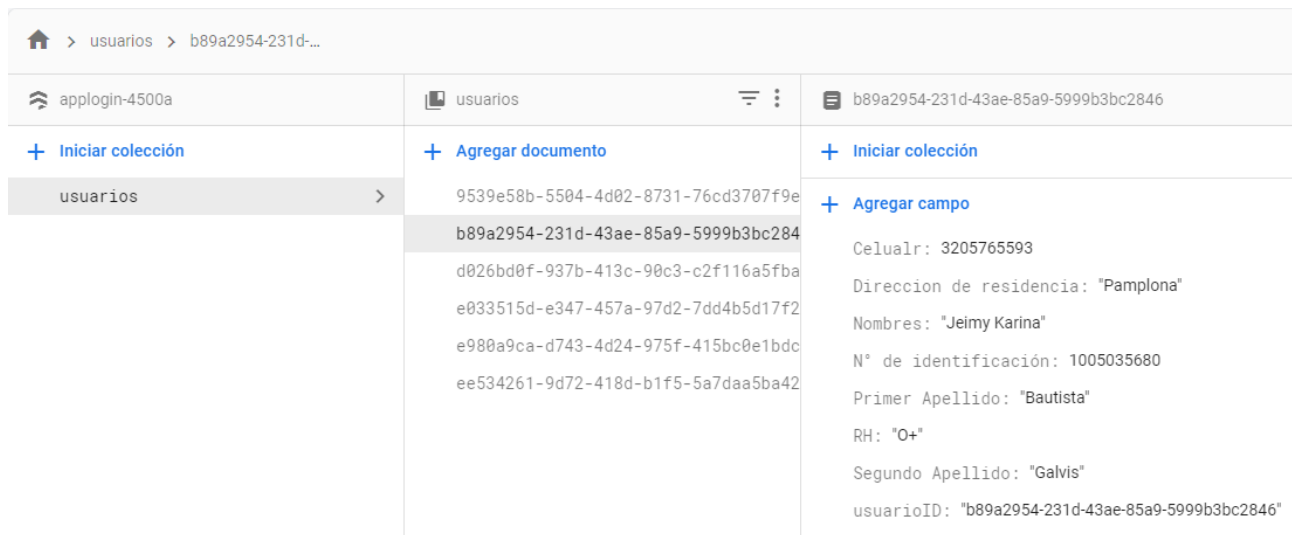


Figura 3.26: Ejemplo de Base de datos de un usuario en Cloud Firestore con sus respectivos datos (Datos no reales solo para uso del ejemplo).

Capítulo 4

RESULTADOS

Ya conociendo la metodología con la cual se abordó el desarrollo de este trabajo, se enseñará los resultados obtenidos abordándolos en tres secciones. Primero, el reconocimiento de los usuarios registrados en la app móvil creada. Segundo, una comparación entre los dos métodos que se usaron para la clasificación de los rostros y tercero, un estudio económico, el cual dará una aproximación muy cercana en términos de dinero al costo de implementar este proyecto.

4.1. Reconocimiento y acceso de datos

En este apartado se muestra la validación de la segunda y tercera screen de la aplicación creada (reconocimiento y registro de usuarios). La información de los usuarios registrados "viajan" en tiempo real a la base de datos en firebase guardándose en apartados únicos creados solamente para este proyecto.

La primera validación se realizó con el creador de este proyecto, en las figuras 4.1 y 4.2 se puede constatar el reconocimiento del rostro y la información guardada de este usuario. Se puede decir que este fue el más sencillo ya que fue el usado para pruebas en toda la realización del proyecto.

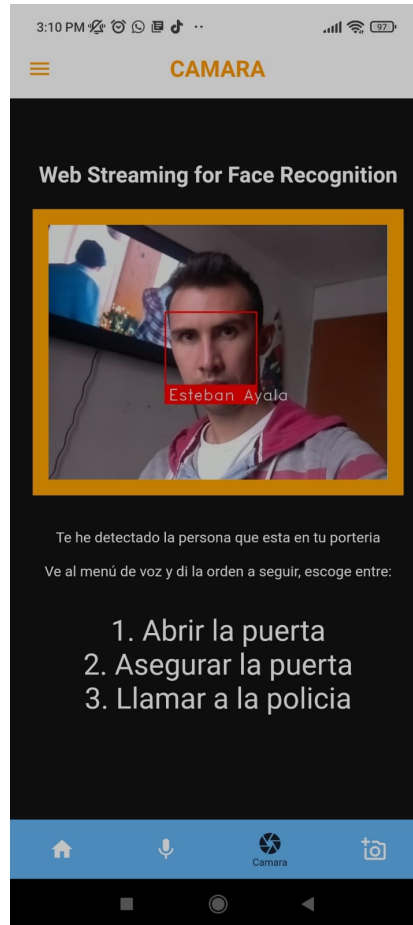


Figura 4.1: Detección de rostro en la app móvil del usuario1(Esteban Ayala).

🏠 > usuarios > 5b3ca2ed-1a34-...		
🔍 applogin-4500a	📁 usuarios	📄 5b3ca2ed-1a34-4d9e-b7dd-547be7a09b12
+ Iniciar colección	+ Agregar documento	+ Iniciar colección
usuarios >	5b3ca2ed-1a34-4d9e-b7dd-547be7a09b1	+ Agregar campo
		Celular: 310280 [redacted] Direccion de residencia: "Carrera 7 a [redacted]" Nombres: "Nelson Esteban" N° de identificación: 101844 [redacted] Primer Apellido: "Ayala" RH: "B+" Segundo Apellido: "Gomez" usuarioID: "5b3ca2ed-1a34-4d9e-b7dd-547be7a09b12"

Figura 4.2: Datos guardados en firestore del usuario1 eliminando parte de ellos por su confidencialidad.

El segundo usuario a identificar fue una mujer adulta para probar el buen funcionamiento del

algoritmo tanto en hombres como en mujeres. Los resultados son los esperados identificando perfectamente al usuario y agregando sus datos a la base de datos.



Figura 4.3: Detección de rostro en la app móvil del usuario2 (Zoley Gómez).

usuarios > faaed30e-9cca...		
applogin-4500a	usuarios	faaed30e-9cca-4a4a-aae3-ecb65e3dfbad
+ Iniciar colección	+ Agregar documento	+ Iniciar colección
usuarios >	5b3ca2ed-1a34-4d9e-b7dd-547be7a09b1	+ Agregar campo
	faaed30e-9cca-4a4a-aae3-ecb65e3dfba	Celular: 313390 [REDACTED]
		Direccion de residencia: "Calle 11 # [REDACTED]"
		Nombres: "Zoley"
		N° de identificación: 60251 [REDACTED]
		Primer Apellido: "Gomez"
		RH: "O+"
		Segundo Apellido: "Quintero"
		usuarioID: "faaed30e-9cca-4a4a-aae3-ecb65e3dfbad"

Figura 4.4: Datos guardados en firestore del usuario2.

Como tercera validación se hizo uso de la imagen de una mujer mas joven en compañía del usuario 1, así se constata de la exactitud del algoritmo. En las figuras 4.5 y 4.6 se observa la identificación de estos usuarios y su base de datos.

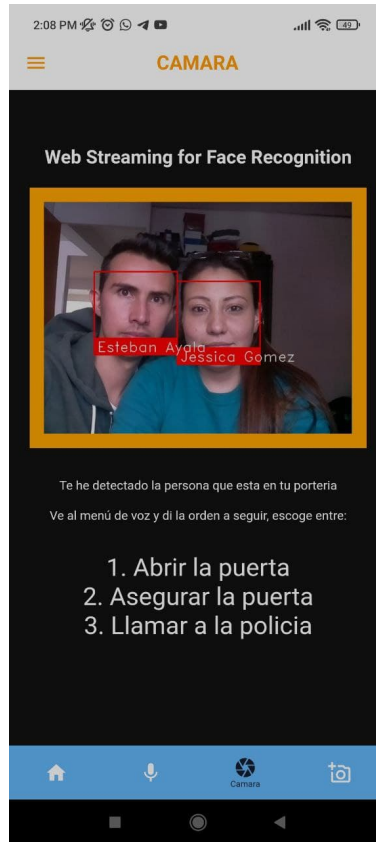


Figura 4.5: Detección de rostro en la app móvil del usuario3 (Jessica Gomez).

usuarios > 6d9ba82b-2b4f-...		
aplogin-4500a	usuarios	6d9ba82b-2b4f-48f8-906c-4c0e15973690
+ Iniciar colección	+ Agregar documento	+ Iniciar colección
usuarios	5b3ca2ed-1a34-4d9e-b7dd-547be7a09b1	+ Agregar campo
	6d9ba82b-2b4f-48f8-906c-4c0e1597369	Celular: 3175251
	faaed30e-9cca-4a4a-aae3-ecb65e3dfba	Dirección de residencia: "Carrera 9 A #"
		Nombres: "Jessica Jaditza"
		N° de identificación: 1094271
		Primer Apellido: "Gomez"
		RH: "A+"
		Segundo Apellido: "Carvajal"
		usuarioID: "6d9ba82b-2b4f-48f8-906c-4c0e15973690"

Figura 4.6: Datos guardados en firestore del usuario3.

Otro usuario validado es una joven en posición de frente, este usuario tiene una particularidad y es que no se realizó pruebas anteriores con ella. Solo al momento de mostrar estos resultados se hizo uso de su imagen y en la figura 4.7 se puede observar el éxito obtenido.

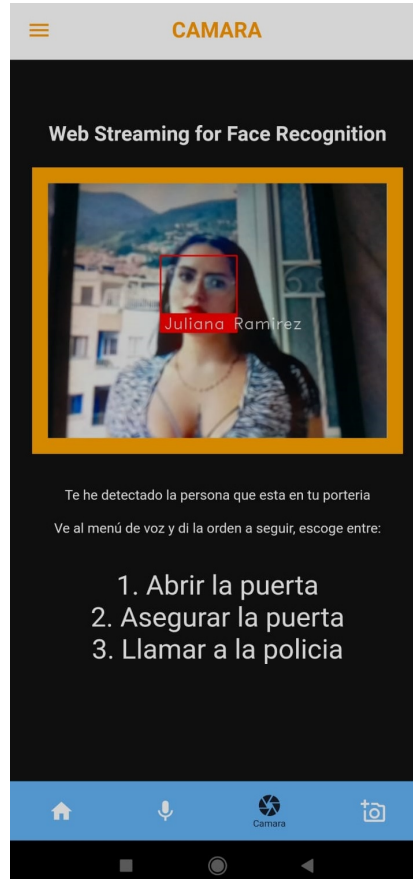


Figura 4.7: Detección de rostro en la app móvil del usuario4 (Juliana Ramirez).

applogin-4500a	usuarios	b4c5a363-4df1-44d2-a7b8-abfe6b4337b9
+ Iniciar colección	+ Agregar documento	+ Iniciar colección
usuarios >	13f836d8-0a8e-4faf-bfc2-065c21c6b1d 5b3ca2ed-1a34-4d9e-b7dd-547be7a09b1 6d9ba82b-2b4f-48f8-906c-4c0e1597369 b4c5a363-4df1-44d2-a7b8-abfe6b4337b c1ee99a3-13f4-49c7-946e-76b2e9a7e92 faaed30e-9cca-4a4a-aae3-ecb65e3dfba	+ Agregar campo Celular: 3123- Direccion de residencia: "calle 7 # 3- Nombres: "Juliana" N° de identificación: 1094- Primer Apellido: "Ramirez" RH: "A+" Segundo Apellido: "Cote" usuarioID: "b4c5a363-4df1-44d2-a7b8-abfe6b4337b9"

Figura 4.8: Datos guardados en firestore del usuario4.

En las figuras 4.9 y 4.11 se puede observar dos usuarios mas utilizados en la validación del proyecto. Primero el usuario 5 el cual es un hombre joven y como usuario 6 un hombre adulto con rasgos faciales muy marcados. Estos usuarios tambien son probados una sola vez. El usuario 5 presentó una pequeña falla en primera instancia ya que no detectaba su rostro de manera clara pero luego se verificó que es un problema de iluminación. Esto brinda la información que con los equipos usados en el proyecto se debe tener una buena iluminación para la cámara.

Se intentó probar en la noche pero presenta dificultades a la hora de la clasificación. De esto se hablará en las conclusiones.

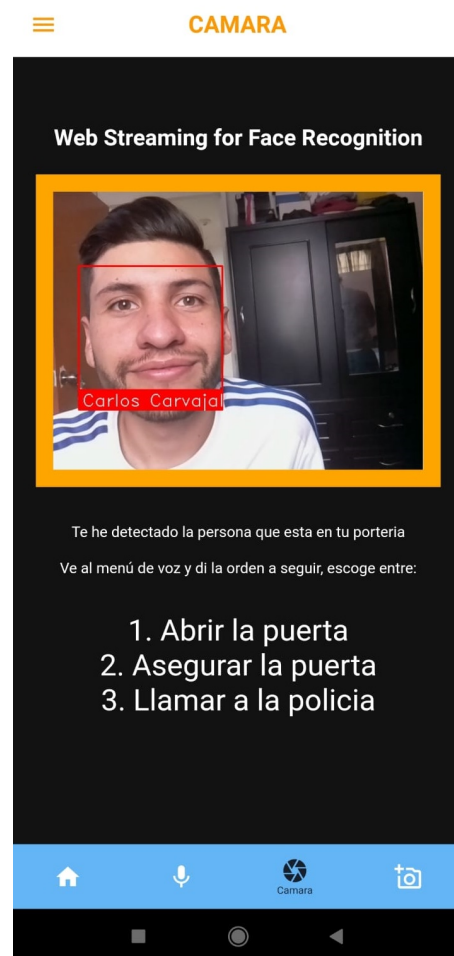


Figura 4.9: Detección de rostro en la app móvil del usuario5 (Carlos Carvajal).

applogin-4500a	usuarios	13f836d8-0a8e-4faf-bfc2-065c21c6b1da
+ Iniciar colección	+ Agregar documento	+ Iniciar colección
usuarios >	04e30ae9-aed6-4089-af6c-490544f2e36	+ Agregar campo
	13f836d8-0a8e-4faf-bfc2-065c21c6b1d	Celular: 320 [REDACTED]
	4f2b5259-ec66-4afc-8c38-5aecaea2549	Dirección de residencia: "calle 9 - [REDACTED]"
	5b3ca2ed-1a34-4d9e-b7dd-547be7a09b1	Nombres: "Carlos Adolfo"
	6d9ba82b-2b4f-48f8-906c-4c0e1597369	N° de identificación: 1096 [REDACTED]
	796ffc6d-5bb5-46a6-b7aa-42941664995	Primer Apellido: "Gomez"
	9504b2d6-f68b-4c99-baa7-416944bb148	RH: "B+"
	aa0676bd-63db-43c4-9f0e-605973abe5d	Segundo Apellido: "carvajal"
	c1ee99a3-13f4-49c7-946e-76b2e9a7e92	usuarioID: "13f836d8-0a8e-4faf-bfc2-065c21c6b1da"
	faaed30e-9cca-4a4a-aae3-ecb65e3dfba	

Figura 4.10: Datos guardados en firestore del usuario5.

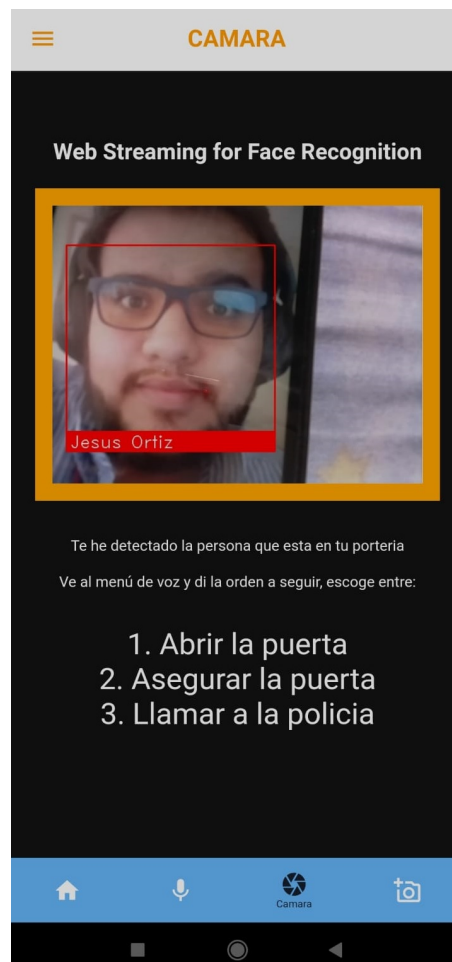


Figura 4.11: Detección de rostro en la app móvil del usuario6 (Jesús Ortiz).

aplogin-4500a	usuarios	c1ee99a3-13f4-49c7-946e-76b2e9a7e92d
+ Iniciar colección	+ Agregar documento	+ Iniciar colección
usuarios >	04e30ae9-aed6-4089-af6c-490544f2e36 4f2b5259-ec66-4afc-8c38-5aecaea2549 5b3ca2ed-1a34-4d9e-b7dd-547be7a09b1 6d9ba82b-2b4f-48f8-906c-4c0e1597369 796ffc6d-5bb5-46a6-b7aa-42941664995 9504b2d6-f68b-4c99-baa7-416944bb148 aa0676bd-63db-43c4-9f0e-605973abe5d c1ee99a3-13f4-49c7-946e-76b2e9a7e92 faaed30e-9cca-4a4a-aae3-ecb65e3dfba	+ Agregar campo Celular: 317 [redacted] Direccion de residencia: "carrera 8 [redacted]" Nombres: "Jesús" N° de identificación: 109 [redacted] Primer Apellido: "ortiz" RH: "O-" Segundo Apellido: "sandoval" usuarioID: "c1ee99a3-13f4-49c7-946e-76b2e9a7e92d"

Figura 4.12: Datos guardados en firestore del usuario6.

Todos los usuarios autorizaron dar su datos para el tratamiento de estos SOLO en la aplicación creada.

4.2. Comparación entre los dos métodos de detección

En esta sección se aborda la comparación de los dos métodos utilizados para la identificación y reconocimiento facial de los usuarios evaluados en la sección anterior. Se comparará el accuracy, la sensibilidad y el recall de los mismos usuarios haciendo uso del primer y segundo método y así poder conclusiones con cifras que lo sustenten.

El primer método a utilizar es el clásico y mas usado en programacion en python el cual es la libreria OPENCV. El segundo método es haciendo uso de un tema que esta en auge en la actualidad como son las redes neuronales, más específicamente las redes neuronales convolucionales que son perfectas para utilizarlas en identificación de rostros.

En la siguiente tabla se puede observar los resultados obtenidos de accuracy de los dos métodos.

	OpenCV(%)	MTCNN(%)
Usuario 1	85.2	91.3
Usuario 2	83.6	92.6
Usuario 3	81.7	92.2
Usuario 4	89.1	94.5
Usuario 5	87.2	93.5
Usuario 6	87.4	94.9

Tabla 4.1 Comparación del parámetro accuracy en los métodos de clasificación utilizados.

En la tabla 4.2 se puede observar los resultados obtenidos de sensibilidad de los dos métodos.

	OpenCV(%)	MTCNN(%)
Usuario 1	81.2	93.3
Usuario 2	83.5	94.2
Usuario 3	86.3	96.7
Usuario 4	89.4	92.7
Usuario 5	85.2	94.8
Usuario 6	83.6	91.2

Tabla 4.2 Comparación de del parámetro sensibilidad en los dos métodos.

Como se pudo observar en las tablas anteriores, el mejor método para la detección de rostros en la app móvil creada son las redes neuronales convolucionales al presentar un mejor accuracy y sensibilidad. Otro factor importante a resaltar es el recall, la métrica de recall o exhaustividad nos va a informar sobre la cantidad que el modelo de machine learning es capaz de identificar. En la

tabla 4.3 se observan estos resultados.

	OpenCV(%)	MTCNN(%)
Usuario 1	81.2	95.3
Usuario 2	82.6	88.5
Usuario 3	87.2	87.2
Usuario 4	88.1	92.9
Usuario 5	88.8	91.9
Usuario 6	82.9	89.4

Tabla 4.3 Comparación del parámetro recall entre métodos usados.

4.3. Estudio económico del proyecto

Como en todo proyecto, y más de este tipo, se generan unos gastos de implementación tanto fijos como mensuales, estos gastos están divididos en dispositivos físicos como en licencias de softwares o rentas de servidores.

En este proyecto específicamente se seleccionaron los dispositivos más acordes para un buen rendimiento pero siempre pensando también en el factor económico. En la tabla 4.3 se observa detalladamente el valor de cada elemento y un total de los gastos fijos para una posterior implementación en una StartUp tecnológica a la cual apunta este trabajo de grado.

Dispositivo	Valor en pesos col.	Valor en dólares
Raspberry Pi 8 GB	520.000	135
Accesorios Raspberry	75.000	19
Módulo de cámara	58.000	15
Diseñador MockUp App	140.000	34
Licencia App Store	387.000	100
Play Store	97.000	25

Tabla 4.3 Gastos fijos en la elaboración del proyecto incluyendo personal contratado e implementos físicos.

Según la tabla anterior, se observa que el dispositivo más costoso es la raspberry Pi. Es entendible que lo sea ya que es el dispositivo que se utilizó como servidor propio y de unico acceso al administrador. Al haber diversas versiones de esta Raspberry Pi con menor capacidad en su memoria RAM, hubiera sido posible aminorar el costo total del proyecto, pero se quedaba corto al momento de realizar el procesamiento de los algoritmos de reconocimiento por lo que se hace necesario hacer uso de un dispositivo con un alta memoria RAM.

Tambien se presentan otros costos que son mensuales los cuales se especifican en la tabla 4.4. Estos gastos son las rentas de uso del servidor creado para poder tener el servicio completo.

Dispositivo	Valor en pesos col.	Valor en dólares
Mantenimiento del Servidor	78.000	20

Tabla 4.4 Gastos mensuales en la elaboración del proyecto los cuales se deben renovar para el buen funcionamiento de este.

Capítulo 5

CONCLUSIONES

Al trabajar con frames de video y su posterior procesamiento se presentan varias dificultades, algunas de estas de gran complejidad. Estas dificultades se deben a las características propias de las imágenes con el agravante de no haber podido obtener la data de la forma deseada en el principio de este trabajo, por lo que se presentan partes que no estaban contempladas.

Con el paso del tiempo y la realización del proyecto, surgieron más detalles, con lo que se hizo preciso utilizar más herramientas en la etapa de reconocimiento facial para hacerla más robusta y poder trabajar mejor.

Luego de haber realizado esta etapa de reconocimiento, el cual se considera que fué la de mas complejidad, se logra realizar un tratamiento de la voz para el apartado de los comandos de voz y que este fuera mucho más limpio haciendo uso de varias librerías dando un resultado 99 porc. veraz. Después de comparar las 6 pruebas empleadas para la validación de los algoritmos, dieron como resultado porcentajes de error bastante buenos, aunque algunos valores se podrían considerar altos, son aceptables en el método de OpenCV ya que precisamente un fin de este proyecto es de-

mostrar la buena utilidad de las redes neuronales en comparación a otros métodos. Lo cual permite concluir, que la aplicación creada le permitirá al usuario tener un excelente método de seguridad en sus viviendas o construcciones.

De lo anterior se puede deducir que el algoritmo es bastante eficiente en algunas imágenes y poco menos en otras, esto debido mayormente a la forma de la adquisición de la data donde las imágenes necesitan de luz natural para que la cámara detecte de forma excelente al usuario, la luz artificial o con poca luz natural influye en el tiempo en que la aplicación demora en reconocer al usuario, esto se logró subsanar en gran medida mediante los filtrados, pero no totalmente. En vista de esto se recomienda realizar la adquisición de las imágenes en un lugar estándar de día para así poder tener mejores resultados.

La Ingeniería en Telecomunicaciones es un programa con extenso campo de aplicación, donde se encuentra la gestión de redes telemáticas, el desarrollo de software, el procesamiento de imágenes o señales, microcontroladores, entre otras. La línea de escogencia para este trabajo fue el desarrollo de software con procesamiento de imágenes, la cual permitió a diferentes usuarios poder tener un sistema de seguridad innovador y poco visto en la ciudad. Al estar el mundo en constante cambio se observa que las edificaciones no solo deben evolucionar en sus formas arquitectónicas sino también en sus sistemas de seguridad e interacción con las diferentes ubicaciones dentro de un apartamento, casa, etc.

Como conclusión final, se puede decir que con el desarrollo de este proyecto se afianzó conocimientos adquiridos en diversas asignaturas del pregrado, a pesar de nunca haber trabajado

anteriormente en el ámbito de las aplicaciones móviles la experiencia fue bastante buena y se deja el proyecto abierto a mejoras.

Capítulo 6

Referencias Bibliograficas

Durán Suárez, J. (2017). Redes neuronales convolucionales en R: Reconocimiento de caracteres escritos a mano.

El Naqa, I., Murphy, M. J. (2015). What is machine learning?. In machine learning in radiation oncology (pp. 3-11). Springer, Cham.

Expósito Gallardo, M. D. C., Ávila Ávila, R. (2008). Aplicaciones de la inteligencia artificial en la Medicina: perspectivas y problemas. Acimed.

Goodfellow, I., Bengio, Y., Courville, A. (2016). Deep learning. MIT press.

Hinestroza Ramírez, D. (2018). El Machine Learning a través de los tiempos, y los aportes a la humanidad (Doctoral dissertation, Universidad Libre Seccional Pereira).

Izaurieta, F., Saavedra, C. (2000). Redes neuronales artificiales. Departamento de Física, Universidad de Concepción Chile.

Jones, Time (2018) Los lenguajes de la IA

Junestrand, S., Passaret, X., VAZQUEZ ALVAREZ, D. A. N. I. E. L. (2004). Domótica y hogar digital. Editorial Paraninfo.

Lastra Fernandez, Emilio (2018) Usos y aplicaciones de la Inteligencia Artificial en el Marketing

Leyva-Vázquez, M., Smarandache, F. (2018). Inteligencia Artificial: retos, perspectivas y papel de la Neutrosología. Infinite Study.

Matich, D. J. (2001). Redes Neuronales: Conceptos básicos y aplicaciones. Universidad Tecnológica Nacional, México, 41.

McCorduck, P., Cfe, C. (2004). Machines who think: A personal inquiry into the history and prospects of artificial intelligence. CRC Press.

Muggleton, S. (2014). Alan Turing and the development of Artificial Intelligence. AI communications, 27(1), 3-10.

Núñez Sánchez-Agustino, F. J. (2016). Diseño de un sistema de reconocimiento automático de matrículas de vehículos mediante una red neuronal convolucional.

Rich, E., Knight, K., Calero, P. A. G., Bodega, F. T. (1994). Inteligencia artificial (Vol. 1). McGraw-Hill.

Roman Victor Introducción al Machine Learning: Una Guía Desde Cero, 2019

SOTO, C., JIMÉNEZ, C. (2011). Aprendizaje supervisado para la discriminación y clasificación difusa. Dyna, 78(169), 26-33.

Tello, J. C., Informáticos, S. (2007). Reconocimiento de patrones y el aprendizaje no supervisado. Universidad de Alcalá, Madrid.

Yan, L. C., Yoshua, B., Geoffrey, H. (2015). Deep learning. nature, 521(7553), 436-444.