

# Universidad del Valle de Guatemala

Facultad de Ingeniería

Departamento de Ciencia de la Computación y Tecnologías de la Información



Análisis de comunicación no verbal en individuos para detección de emociones humanas  
mediante técnicas de inteligencia artificial y aprendizaje de máquina

Trabajo de graduación presentado por

José Fernando Flores Calderón

para optar al grado académico de Licenciado en Ingeniería en Ciencias de la  
Computación y Tecnologías de la Información

GUATEMALA

2018



Análisis de comunicación no verbal en individuos para detección de emociones humanas mediante técnicas de inteligencia artificial y aprendizaje de máquina

# Universidad del Valle de Guatemala

Facultad de Ingeniería

Departamento de Ciencia de la Computación y Tecnologías de la Información



Análisis de comunicación no verbal en individuos para detección de emociones humanas  
mediante técnicas de inteligencia artificial y aprendizaje de máquina

Trabajo de graduación presentado por


José Fernando Flores Calderón

para optar al grado académico de Licenciado en Ingeniería en Ciencias de la  
Computación y Tecnologías de la Información

GUATEMALA

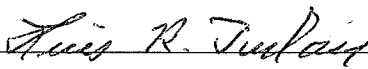
2018

Vo.Bo.:

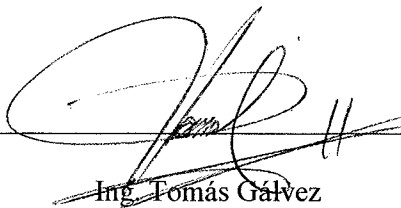
(f)  \_\_\_\_\_

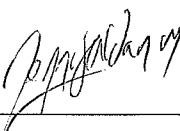
Ing. Dennis Aldana

Tribunal Examinador:

(f)  \_\_\_\_\_

Ing. Luis Furlán

(f)  \_\_\_\_\_  
Ing. Tomás Gálvez

(f)  \_\_\_\_\_

Ing. Dennis Aldana

Fecha de Aprobación: Guatemala, 3 de diciembre de 2018

## PREFACIO

El problema al que nos enfrentamos en este proyecto, comenzó al saber que no existe una forma de detección de emociones humanas que incluya el análisis del lenguaje corporal de un ser humano, a través de mecanismos automatizados de computación. Lo que sí existe es la detección de las emociones a través de la cara, pero no un análisis holístico que incluya ambos gestos de la cara y cuerpo.

Por este motivo, se inició este proyecto con la idea de detectar la emoción de una persona en un instante, a través del análisis del cuerpo y los gestos de la cara. El proyecto se realizó de la siguiente forma: Se plantearon 4 módulos principales, los cuales fueron: etiquetado de imágenes (validación de etiquetado), estimación de pose, normalización y discretización de los datos e implementación del algoritmo de aprendizaje supervisado (machine learning).

Para validar que el etiquetado de imágenes estuviera correcto, se procedió a estudiar el lenguaje corporal y las expresiones faciales y además se verificaron las imágenes, con la ayuda de la licenciada Lorena Gómez, psicóloga de una institución educativa.

Se construyó un compilado de 400 imágenes para el entrenamiento del algoritmo, etiquetando cada imagen con las emociones humanas básicas. Posteriormente y por medio del algoritmo de regresión logística, el cual es parte del ámbito de machine learning, se procedió a la resolución del problema a través de la predicción de las emociones mediante imágenes de personas.

# ÍNDICE

Prefacio	ix
Lista de figuras	xii
Resumen	xiv
I. Introducción	1
II. Objetivos	2
A. Objetivo general.....	2
B. Objetivos específicos.....	2
III. Justificación	3
IV. Marco teórico	4
A. Emociones.....	6
1. Entender las emociones a través de la cara.....	6
2. Las seis emociones básicas.....	10
3. Cómo funciona el lenguaje corporal.....	13
4. Gestos Universales.....	15
5. Cómo leer y analizar gestos corporales.....	15
6. El poder de la sonrisa.....	17
7. El uso de los brazos.....	17
8. El uso de los pulgares.....	19
9. Gestos de procrastinación.....	19
10. Gestos de evaluación.....	20
11. La importancia de las piernas en el lenguaje corporal.....	20
12. Posiciones básicas de la cabeza.....	22
B. Aprendizaje de máquina.....	25
1. Algoritmo de regresión logística.....	27
2. Supuestos del algoritmo de regresión logística.....	28

3. Estimación Humana de Poses.....	29
4. Algoritmo PoseNet.....	30
5. Funcionamiento del algoritmo PoseNet.....	31
V. Marco metodológico	33
A. Recopilación de imágenes y etiquetado.....	34
B. Elección del modelo de clasificación de emociones.....	36
C. Obtención de los datos.....	37
D. Normalización de los datos.....	38
E. Discretización de los datos.....	39
VI. Resultados	42
VII. Análisis de resultados	48
VIII. Conclusiones	51
IX. Recomendaciones	52
X. Referencias bibliográficas	53
XI. Anexos	55



## LISTA DE FIGURAS

1. Posición de atención.....	21
2. Posición piernas abiertas.....	21
3. Posición pie adelante.....	22
4. Posición piernas cruzadas.....	22
5. Posición cabeza arriba.....	23
6. Posición cabeza hacia un lado.....	23
7. Posición cabeza abajo.....	23
8. Posición hombros encogidos.....	24
9. Posición manos sobre la cintura.....	24
10. “Postura del vaquero”.....	25
11. Función sigmoide .....	27
12. Ecuación de regresión logística .....	27
13. Gráfica de la función sigmoide.....	28
14. Puntos clave de una estimación de pose.....	30
15. Tabla de puntos clave .....	31
16. Resultado del análisis de imagen.....	32
17. Emoción de sorpresa.....	34
18. Emoción de miedo.....	35
19. Emoción de enojo.....	35
20. Emoción de felicidad.....	35
21. Emoción de tristeza.....	35
22. Emoción de disgusto.....	36
23. Salida de la aplicación con una imagen etiquetada “feliz”.....	37
24. Datos sin normalizar y normalizados.....	38
25. Malla para discretización.....	39
26. Datos finales discretizados.....	40
27. Gráfica de data clasificada.....	42
28. Resultado de matriz de predicciones.....	42
29. Resultado del modelo (score).....	43
30. Regresión logística.....	43
31. Tabla de matriz de confusión.....	43

32. Precisión del modelo.....	44
33. Tabla de reporte de clasificación.....	44
34. Predicción de nuevos valores.....	44
35. Data final extraída y normalizada (Imagen etiquetada cómo alegría)	45
36. Data final extraída y normalizada (Imagen etiquetada como miedo)	46
37. Data final extraída y normalizada (Imagen etiquetada como triste)	46
38. Data final extraída y normalizada (Imagen etiquetada como enojo)	46
39. Data final extraída, normalizada (Imagen etiquetada como sorpresa)	47
40. Data final extraída, normalizada (Imagen etiquetada como disgusto)	47
41. Predicciones obtenidas.....	48
42. Persona feliz clasificada como enojada.....	49
43. Persona enojada clasificada como feliz.....	49
44. Persona triste clasificada como enojada.....	49
45. Estimación de pose, con imagen personal.....	55
46. Código para estimación de pose.....	55

## RESUMEN

El lenguaje corporal y el análisis de las emociones humanas juegan un papel importante en nuestra sociedad. El análisis correcto de las emociones en una persona nos provee ideas de actitudes e intenciones. En los últimos años se han empezado a desarrollar librerías de análisis de las poses del cuerpo, las cuales logran detectar las articulaciones del cuerpo humano con bastante precisión. Esta investigación se enfoca en varias posibles aplicaciones de estas herramientas, las cuales pueden ser la detección de movimientos en un videojuego de realidad aumentada, o usos en el ámbito deportivo. En este trabajo se desarrolla la detección de emociones humanas a través del análisis de las poses de las personas. Se utilizan las emociones básicas, que son las más generales posibles. Por medio de un algoritmo de aprendizaje de máquina y data de imágenes recopiladas (que fueron etiquetadas), se construye un modelo de regresión logística para poder obtener la emoción que predomina en una imagen no etiquetada.

# I. INTRODUCCIÓN

La detección de las emociones en los seres humanos es una tarea que nos ayuda a entender mejor a las otras personas en el ambiente en el que nos desenvolvemos. Este trabajo se centra en el análisis de dichas emociones y la implementación de una herramienta que logre clasificar las emociones básicas de una persona. Para ello se utilizó un algoritmo de estimación de poses y un algoritmo de aprendizaje de máquina y estadística, llamado algoritmo de regresión logística. Se desarrolló una base de datos de imágenes de personas de 400 imágenes, de las 6 emociones básicas del ser humano. A pesar de que los algoritmos de estimación de pose interpretan videos, solo se trabajó con imágenes estáticas.

Respecto a las emociones, se busca encontrar las predominantes generalizadas en un individuo. Se detectan las emociones básicas, que incluyen alegría, enojo, miedo, tristeza, sorpresa y disgusto. En este trabajo no nos enfocamos en encontrar emociones más específicas, derivativas de las emociones básicas. Ejemplos de emociones que no se considerarán son: Impaciencia, nerviosismo, excitación, etc.

También es importante considerar que analizaremos la “comunicación no verbal”. Al decir comunicación no verbal, nos referimos a los gestos corporales de la cabeza, cara y los miembros del cuerpo tales como brazos, torso y hombros. La comunicación verbal, también forma parte de la comunicación, pero no se trabajó con ella.

## II. OBJETIVOS

### A. Objetivo general

- Detectar las emociones de una persona en un instante determinado, visualizando el cambio de emociones cuando la persona cambia de postura o realiza un movimiento; mediante la construcción de una solución de software con la aplicación de teoría de inteligencia artificial y aprendizaje de máquina.

### B. Objetivos específicos

- Desarrollar una base de datos de 400 imágenes de personas que sean claras y visibles y que muestren varios gestos y posiciones corporales, además de gestos faciales, clasificando cada imagen según la emoción predominante en la imagen, en alguna de las siguientes cuatro emociones básicas: alegría, tristeza, enojo/rabia y sorpresa.
- Implementar y entrenar un algoritmo de aprendizaje supervisado, con el fin de lograr reconocer las emociones básicas en las imágenes y mediante la utilización del compilado de imágenes etiquetadas.

### III. JUSTIFICACIÓN

Posterior a la realización de este proyecto, existen numerosas formas de utilizarlo como base para otros proyectos e ideas. Al sentar las bases para encontrar información útil del movimiento y gesticulación de las personas, se puede ver que existen aplicaciones de estos conocimientos en muchos campos.

Entre las diversas posibles aplicaciones a futuro, están: la seguridad en lugares públicos mediante detección de comportamientos anormales en las personas, el aprendizaje de la técnica utilizada en movimientos corporales en deportes de raqueta, el lanzamiento de jabalina, el salto de pértiga, y cualquier otro deporte que involucre un movimiento que necesite una técnica específica.

En un ambiente social también tiene aplicación. Por ejemplo, para un restaurante la información de las emociones en las personas tiene utilidad, pues puede permitir medir el estado de aceptación del restaurante en términos de la calidad del ambiente, la música, la comida, el servicio, etc. Al tener información de cómo reacciona la gente como una forma de servicio del restaurante en todos los aspectos mencionados anteriormente, el restaurante tiene la opción de variar dichos aspectos y considerar si funcionan mejor que con los anteriores.

Por otra parte, en el ámbito educativo la aplicación puede ser de utilidad en varias situaciones. Por ejemplo, un catedrático que quiere realizar una medición de la atención e interés en su clase y analizar cómo reaccionan los alumnos ante la forma de impartir el curso. Con la información de las emociones de sus alumnos, puede hacer modificaciones a su forma de impartir la cátedra, y hacer los ajustes necesarios para impartir de mejor manera la clase.

## IV. MARCO TEÓRICO

A continuación se tratarán diversos temas con los que se fundamenta la investigación de este proyecto, tratando temas como inteligencia artificial, aprendizaje de máquina o aprendizaje de máquina y psicología del lenguaje corporal.

El análisis del cuerpo humano y la comunicación no verbal que existe en el proceso comunicativo entre seres humanos es un campo importante de la psicología. La comunicación no verbal permite encontrar información sobre sentimientos, emociones y pensamientos de un individuo, sin que dicho individuo emplee la comunicación explícita mediante el lenguaje verbal. Existen diversos estudios en el campo de la psicología, que estudian el lenguaje no verbal y analizan el significado de este tipo de lenguaje. Para los seres humanos, la lectura e interpretación del lenguaje no verbal nos provee una idea sobre el estado emocional de otras personas. La inteligencia social nos permite asociar los gestos corporales, posturas y movimientos con estados emocionales y pensamientos.

La comunicación es la forma en que los seres humanos transmiten información entre sí, con el objetivo de influir en los demás, y expresar sentimientos e ideas. En el proceso de la comunicación existen dos partes, un emisor y un receptor. El emisor es la persona o ente que emite el mensaje, pudiendo ser una persona o no. Por otro lado, el receptor es el encargado de recibir el mensaje. Hay dos tipos de comunicación entre personas, verbal y no verbal. Esta última es el tipo de comunicación en la que el habla no es requerido.

Los seres humanos podemos expresarnos mediante la utilización de gestos, movimientos y acciones con el cuerpo. Esta forma de expresión se denomina como lenguaje corporal no verbal, en donde no es necesario utilizar el lenguaje verbal para la comunicación de ideas y pensamientos. Las personas pueden mentir, pero la comunicación no verbal inconscientemente habla la verdad. El lenguaje corporal juega un papel crucial en los mensajes que damos y recibimos de otra persona. Este concepto incluye expresiones faciales, contacto visual, posturas, apariencias, sonrisa y comportamiento. También se conoce el lenguaje corporal como “kinética”.

En este trabajo, el enfoque se dio en los gestos corporales, la posición de las partes del cuerpo y las expresiones faciales. Se debe mencionar que esto último es relevante, puesto que la expresión en la cara está directamente relacionada con la expresión en el cuerpo. Un gesto se refiere a cualquier movimiento significativo que realice la persona y el uso deliberado de dicho movimiento como una expresión o sentimiento por medio de un miembro del cuerpo. Como ejemplos se pueden mencionar: el movimiento de la cabeza (cabeceo), la sonrisa, el movimiento de los labios, la colocación del brazo en el hombro, tocarse la cabeza, señalar en una dirección, entre otros.

En lo que respecta a la teoría de emociones básicas, en el libro “*Unmasking the face*” (Ekman, 2003), se mencionan que las 6 emociones básicas que el ser humano experimenta son: alegría, tristeza, enojo, sorpresa, miedo y disgusto. A partir de estas emociones, se derivan el resto de emociones, las cuales según el autor, son combinaciones de estas emociones básicas.

Otro de los temas computacionales que se utilizarán en este trabajo, es la inteligencia artificial. Se define de la siguiente manera: el estudio de agentes que toman percepciones de su entorno, para llevar a cabo determinadas acciones. (Russell, S. J., Norvig, P., & Canny, J, 2003). Cuando se habla de inteligencia artificial, se debe mencionar el concepto de racionalidad; Se define como racional un sistema que hace lo “correcto”, en base del conocimiento que tiene. Se tratarán los temas utilizados sobre estas áreas de estudio más adelante. (Russell, S. J., Norvig, P., & Canny, J, 2003)

En inteligencia artificial, un agente es una entidad que tiene capacidad de percibir su ambiente o entorno, y actuar de manera correcta. Cada agente se encarga de realizar funciones específicas, las cuales a su vez estructuran las secuencias de las percepciones para la implementación de acciones. Existen diferentes formas de representar estas funciones (las que el agente realiza), tales como agentes reactivos, planificadores condicionales, redes neuronales y sistemas teóricos para las decisiones, entre otros conceptos.

Otro de los temas que se aplicarán en este trabajo es aprendizaje de máquina o aprendizaje de máquina. Al hablar sobre este campo, se debe hacer notar que es un subtema de inteligencia artificial. La idea principal de aprendizaje de máquina es que mediante algoritmos pueda ser posible encontrar conclusiones e información relevante, tomando como punto de partida un conjunto de datos. A esta idea se debe



agregar que el objetivo de aprendizaje de máquina, a través de algoritmos, el ser humano no tenga que formar parte de la solución de los problemas directamente.

El aprendizaje supervisado es un tipo de algoritmo del campo de estudio de aprendizaje de máquina. El aprendizaje se basa en la idea de tomar las percepciones y utilizarla para mejorar la habilidad del agente, y actuar en el momento dando una solución. El aprendizaje supervisado toma entradas y salidas para aprender y predecir futuros resultados.

En lo que respecta a las soluciones anteriormente realizadas en este campo, se puede mencionar el trabajo realizado por ingenieros de la universidad Carnegie Mellon, quienes han desarrollado un software capaz de identificar los miembros del cuerpo, tales como brazos, piernas, cabeza, además otras articulaciones del cuerpo.

## A. Emociones

### 1. Entender las emociones a través de la cara

Entender las expresiones faciales nos sirve a los humanos para entender las emociones en otras personas e incluso en nosotros mismos. Existen diferentes maneras en que se puede expresar alguien con sus expresiones faciales. Existen variaciones entre la forma en que las personas muestran una misma emoción a través de su cara. Entre los motivos por los que cada quien expresa una misma emoción de forma diferente, están: el condicionamiento social con el que crecimos y de donde aprendimos a detectar cada emoción, la familia, y principalmente los padres y cualquier familiar con el que convivimos en la infancia. (Ekman, 2003)

Las emociones básicas son las emociones generales de las cuales se desprenden todas las emociones específicas. Estas son: alegría, sorpresa, miedo, enojo, tristeza y asco. Cualquier emoción específica es posible relacionarla a una combinación de estas emociones básicas. La cara provee tres tipos de señales, de tipo estático, lento o rápido. Las señales de tipo estático pueden ser el tono de la piel o la forma de la cara; Las señales lentas puede ser las arrugas y las rápidas el movimiento de los párpados u ojos.

Al analizar las emociones según lo que se observa en la cara, se debe mencionar el concepto de “microexpresiones”. La mayoría de expresiones faciales en donde se expresa una emoción, suceden rápidamente. Pueden durar fracciones de segundo en ser expresadas en la cara y son lo que se le conoce como “microexpresiones”. Este

tipo de expresiones puede revelar la emoción de la persona en un instante. Esto es contrario a una macro-expresión que es un tipo de expresión facial que dura más tiempo, entre 2 a 3 segundos, dependiendo de la intensidad de la emoción. Una microexpresión puede ser no comprendida o ni siquiera comunicada puesto que suceden demasiado rápido, por lo que pueden existir malas interpretaciones. Cabe mencionar que es común ver a personas expresar emociones de forma no genuina, es decir, que no están sintiendo esa emoción de la forma que lo proyectan, sino de una manera exagerada que no se apega a los sentimientos que experimentan. A este tipo de expresiones se le denominan expresiones simuladas y se deben tener en cuenta al analizar lo que comunica la cara de un individuo. (Ekman, 2003)

Los mensajes emocionales son comunicados por el tono de voz, movimientos de las extremidades como pies y brazos, postura, gestos corporales y expresiones faciales. La cara es una de las fuentes de información más precisas sobre las emociones. La cara es un claro y preciso sistema de señales para la expresión emocional. Ha sido probado que la gente se concentra más en el canal visual que otros tipos de fuente de información y la atención en la comunicación entre dos personas está más orientada a la cara que a los movimientos corporales (Pease, 2017). El canal auditivo también transmite información y más que todo representa una puja por la atención de la otra persona, lo que puede provocar que se pierda el enfoque en la cara para detección de emociones. Sin embargo, las expresiones faciales tienen más importancia que las palabras para transmitir información. (Pease, 2017)

La cara, además de comunicar señales (como rápidas y estáticas), comunica mensajes acerca de emoción, estado emocional, actitud, inteligencia, atraktividad y otras características. Los sentimientos que se analizaron en esta investigación son los sentimientos transitorios, como enojo, sorpresa y alegría. Al referirnos a emociones, realmente nos referimos a este tipo de sentimientos. Estos son transmitidos por las señales rápidas emitidas por la cara. La forma principal para detectar cuando ocurren este tipo de sentimientos, es analizando la cara y notando contracciones de los músculos de la cara que provocan cambios visibles en el aspecto de la misma. Como ejemplos de este tipo de cambios, están: el movimiento o contracción de los labios barbilla o cejas, la aparición y desaparición de arrugas (en la frente por ejemplo), cambios en la forma de los ojos, pómulos, barbilla o nariz. (Ekman, 2003)

A través de diversos estudios se ha comprobado que se puede obtener precisas estimaciones de las emociones actuales de un individuo mediante el análisis de las

señales rápidas emitidas por la cara. Son este tipo de señales las que nos proveen de una idea acertada de las emociones, mientras que las señales de tipo lento y estáticas han sido probadas que no transmiten mensajes emocionales (Pease, 2017). A pesar de esto, se debe tener en cuenta que las señales lentas y estáticas pueden afectar la estimación final de las emociones. Si por ejemplo, la cara de una persona adulta se observa con arrugas en la frente, podría asociarse erróneamente al enojo o rabia, cuando la persona en realidad está completamente con una emoción diferente.

Es importante diferenciar los conceptos de emoción, sentimiento, estado anímico y actitud. Un sentimiento (o sensación) es causada por la reacción corporal hacia algún evento o circunstancia. Una emoción es la asociación que se le da a dicho sentimiento después que esta ocurre. Por otra parte, el estado anímico es diferente a las emociones ya que los sentimientos involucrados duran un período de tiempo más largo. Por ejemplo, un sentimiento de tristeza que dura unos minutos o incluso horas, se le llama emoción. Pero si la tristeza se prolonga por días, esto se convierte en un estado anímico. En esta investigación el enfoque estará en las emociones, pero se debe aclarar que es posible que al analizar la emoción de una persona en un instante, se confunda con una emoción que está sintiendo la persona por el evento o circunstancia en la que está, aunque realmente está en un estado anímico desde hace varios días y por otros motivos. (Pease, 2017)

También existen los mensajes de tipo emblemático emitidos por los gestos faciales. Estos son parte de las señales rápidas. Un mensaje de tipo emblemático es aquel en el que el gesto expresado sustituye un mensaje verbal. (Ekman, 2003) Por ejemplo, si se levantan las cejas, la emoción asociada podría ser sorpresa, pero si solo se levantan las cejas dejando estático los parpados, ojos y demás partes de la cara, esto lo interpretamos como interrogación.

El fruncido del ojo lo interpretamos como “sí” o como “por supuesto”. Estos mensajes emblemáticos son aprendidos desde la infancia y todas las personas comprenden el significado. Un mensaje de este tipo si fuese analizado fuera de contexto o solo por el movimiento específico de las partes de la cara que cambian, se asociaría a una emoción diferente que no está relacionada con lo que se está expresando. (Ekman, 2003)

Por otro lado, las expresiones faciales pueden ser controladas o no controladas. Una expresión puede ser voluntaria o involuntaria. Esto significa que, si una expresión es controlada, no es una expresión real de la emoción actual de la persona.

Cuando una expresión es involuntaria, no se necesita pensar en cómo modular los músculos de la cara para expresar la emoción, simplemente así es la reacción natural. Es posible que exista una mezcla de una expresión controlada y una no controlada. Esto sucede cuando se está experimentando una emoción, pero al mismo tiempo conscientemente se regula una parte de la cara para parecer de cierta forma o transmitir algo ligeramente diferente.

Existen motivos por los cuales alguien quiera controlar las expresiones faciales. Hay normas y convenciones sociales de lo que está permitido mostrar en la cara. Por ejemplo, una norma aprendida desde la infancia, en donde la familia le instruye al niño a no llorar en frente de otras personas. Cuando el niño esté en una situación en donde llora, el hábito aprendido a través de su familia puede causar que modifique ciertos gestos para disimular lo que siente. También pueden estar controladas por estar en un ambiente laboral o estudiantil. En un trabajo de ventas o de diplomacia, existe una gran importancia en que la persona controle sus gestos faciales y muestre los apropiados para la situación, puesto que representa una ventaja hacerlo a que si mostraran las expresiones reales. (Ekman, 2003)

Existen otras emociones que también son comunicadas por la cara, pero estas no han sido comprobadas ni establecidas firmemente. A esto se le debe agregar que al observar una fotografía de la cara de una persona, esta puede comunicar más de una emoción al mismo tiempo. La mezcla de estas 6 emociones abarca la gran mayoría de las posibles expresiones faciales.

Otro punto necesario de mencionar, es si las expresiones faciales son universales, es decir, se muestran de la misma forma en todas las diferentes culturas del mundo. Según la teoría de Charles Darwin, las expresiones faciales que muestran emoción son universales, biológicamente determinadas y son el resultado de la evolución del hombre. A partir de esta primera teoría, se han realizado investigaciones para comprobar la veracidad de la idea. En un estudio realizado por Paul Ekman, se mostró un filme de película a dos hombres, uno de Japón y el otro de Estados Unidos. (Ekman,2003) Al finalizar el estudio se concluyó que ambos reaccionaron de la misma manera, con muy poca diferencia entre los gestos que mostraron. Por ende, la teoría de Darwin se mantiene y ha quedado evidenciado por este y más estudios que las expresiones faciales de la emoción si son universales y no dependen de lenguaje, cultura o condicionamiento social. (Darwin, 1872)

## 2. Las seis emociones básicas

### a. Sorpresa:

La emoción de sorpresa es la emoción que menos tiempo dura. Si la persona tiene tiempo para pensar en la situación y sobre si está sorprendido o no, ya no lo está. Es una emoción que no perdura, rápidamente se transiciona hacia otra emoción, comúnmente miedo o alegría. Es generada por lo inesperado y lo imprevisto. En caso de que la persona tenga tiempo para anticipar lo inesperado, la emoción de sorpresa ya no ocurrirá. La sorpresa dura solo el tiempo en el que el individuo se tarda en racionalizar lo ocurrido. Si la sorpresa es de agrado para la persona, la emoción que tomará el lugar ahora es alegría. Si por otro lado la sorpresa no es de agrado, la persona puede transicionar al miedo, disgusto o enojo. Debido a que la sorpresa sucede rápidamente, en la cara de la persona comúnmente se observará una mezcla de sorpresa con otra emoción. (Ekman, 2003)

La sorpresa varía en como se observa en la cara de la persona, dependiendo de la intensidad de la emoción, desde leve hasta extremo. Existe una emoción que normalmente se asocia a la sorpresa: el asombro. Esta emoción es causada por un cambio extremo de estimulación; Por ejemplo, el fuerte sonido de un explosivo. (Ekman, 2003)

Los gestos faciales de la sorpresa se muestran en tres partes de la cara: primero los ojos, los cuales son abiertos ampliamente. Segundo las cejas, las cuales se elevan y tercero la mandíbula cae separando los labios. Los ojos son abiertos con los párpados bajos relajados y los párpados bajos elevados. Al momento en el que las cejas se elevan, se producen arrugas horizontales en la frente, aunque no todos los adultos las muestran y muchos niños pequeños tampoco.

### b. Miedo:

Esta emoción se genera por la idea de que existirá daño físico o psicológico a la persona. La emoción es experimentada en un ser humano en varios niveles. Puede ser desde un miedo a algo menor como el rechazo verbal o algo mayor como el dolor físico. Como seres humanos sobrevivimos, debido a nuestra capacidad de evitar daño y escapar de situaciones que nos ponen en peligro. El miedo es aquella emoción que se experimenta previamente a la ocurrencia del evento, que puede suceder o no. Dicha anticipación al evento puede suceder desde minutos hasta semanas antes a la ocurrencia en sí. El miedo también se puede experimentar ya cuando ha sucedido un

evento que nos pone en peligro física o emocionalmente

A diferencia de la emoción de la sorpresa, el miedo es una experiencia negativa en todas sus posibles facetas. En la emoción de sorpresa, esta emoción puede experimentarse de forma positiva o negativa, pero en el miedo, cualquier gama por muy sutil que sea se experimentará ligeramente cómo una sensación negativa. (Ekman, 2003)

Es importante considerar en esta emoción, que si analizamos a una persona estática y presenta lenguaje corporal que da indicios de miedo, dicha emoción puede no estar siendo generada por algo del ambiente en el momento preciso, sino que la persona puede haber experimentado la emoción un momento diferente y presentar lenguaje corporal de miedo. Otra característica del miedo es la duración. El miedo suele ser una de las emociones que más persiste en el tiempo; otras emociones como la sorpresa duran poco, instantes a veces, mientras que el miedo puede durar horas e incluso más.

También es importante saber que el miedo puede ser antecedido por otra emoción, como sorpresa o enojo. Esta otra emoción incluso se puede manifestar posteriormente. Respecto a la apariencia física de la persona cuando experimenta miedo, en la cara las cejas están ligeramente elevadas y juntas, los ojos están juntos, los párpados están tensos y los labios se retraen. (Ekman, 2003)

b. Felicidad:

Esta es la emoción que se siente bien. Es una emoción positiva, en comparación con el miedo, disgusto y tristeza, que son emociones negativas mayoritariamente. La felicidad varía en intensidad también; se puede estar ligeramente feliz o experimentar alegría o éxtasis. Es una emoción que puede ser audible, debido a que está directamente relacionada con la risa. Existen diferentes variantes, pero puede ser demostrada mediante una simple sonrisa, o mediante carcajadas intensas. (Ekman, 2003)

A pesar de esto, hay que tener en consideración que la sonrisa que comúnmente es una señal de felicidad, puede también ser utilizada para esconder otras emociones. Por ejemplo, si la persona experimenta miedo, una sonrisa puede ser usada para enmascarar el miedo y pretender que se está feliz.

Para detectar felicidad en una persona, se debe observar la cara y buscar las siguientes características: los extremos de los labios se retraen para atrás y para arriba, los labios pueden no estar partidos, mostrando los dientes o no, una arruga se

forma desde la nariz para los extremos de la boca, las mejillas se levantan, los párpados tienen arrugas en la parte inferior y pueden estar levantados, pero no tensos. (Ekman, 2003)

c. Tristeza:

Esta es la emoción silenciosa. El sufrimiento en esta emoción no tiene sonidos. Puede ser generada por varios motivos, desde la pérdida de un ser querido, hasta el rechazo de una empresa ante una oferta laboral. Esta emoción no dura poco, sino que comúnmente se alarga por periodos de minutos u horas a días. La persona que experimenta tristeza no está en modo activo, sino pasivo. Es una emoción negativa, así como el miedo o angustia. (Ekman, 2003) Existen diferentes niveles de tristeza, que pueden variar desde sentirse melancólico, hasta tristeza profunda o depresión. La expresión de la tristeza es silenciosa, menos turbulenta, menos vocal que la emoción de angustia. Respecto a la detección en el lenguaje corporal de una persona, esta emoción es una de las más complicadas de detectar en la cara, puesto que la tristeza no emite ningún gesto claramente visible, sino que se transmite de forma sutil. A pesar de esto, existen ciertas características que se pueden observar en la cara del individuo para detectarla. En la cara, los extremos de los párpados superiores están elevados, y el párpado inferior puede estar levantado; los extremos de los labios están inclinados hacia abajo y los labios parecen temblar ligeramente. (Ekman, 2003)

d. Enojo:

Esta emoción es la más peligrosa de todas, ya que a partir del enojo las personas pueden incurrir en practicar la violencia, verbal o física. Puede ser producida de diferentes formas. Una forma es por la frustración ante una situación que no ocurrió como la persona esperaba. Otra forma es mediante el daño emocional producido cuando una persona es atacada de manera psicológica, a través de un insulto o rechazo. También puede ocurrir que el daño está siendo causado a alguien más y al observar dicha acción que va en contra de nuestras morales, sentimos enojo. (Ekman, 2003)

El enojo es muy visible en el lenguaje corporal, y es posible detectarlo. Al estar en este estado, la presión sanguínea aumenta, la cara puede tornarse roja y las venas en la cara y cuello pueden aparecer. Por otra parte, también existen cambios en la respiración, el cuerpo puede tornarse tenso. Esto puede o no estar acompañado de un ataque físico o verbal, si el sentimiento en el cuerpo es demasiado intenso. La intensidad varía desde una ligera irritación, hasta la furia. (Ekman, 2003)

En la cara, los gestos que están relacionados con la emoción son: las cejas están inclinadas hacia abajo y se juntan en el ceño, los párpados están tensos y la mirada está fija fuertemente. Además, los labios están presionados entre sí. (Ekman, 2003)

e. Asco o disgusto:

Esta emoción es expresada mediante un sentimiento de aversión y repugnancia. Podemos experimentar asco al imaginar algo desagradable o al estarlo experimentando en el momento, como por ejemplo comer un alimento expirado. La emoción puede estar causada por cualquier estímulo que recibimos en cualquiera de los 5 sentidos. Desde el olfato para mostrar repugnancia por un olor desagradable, o el sonido de una canción que no apreciamos. También pueden generar asco, no solo los estímulos a los cinco sentidos, sino ideas y percepciones de otras personas (Ekman, 2003) El asco involucra alejarse o evitar la causa de la emoción. La meta al experimentar esta emoción es retirarse del objeto o estímulo que causa la emoción. La intensidad de esta emoción puede variar, desde un ligero disgusto por algo, hasta un disgusto que provoque vómito o náusea. En esta emoción, la nariz y la boca son las principales partes de la cara que en la que se expresa la emoción. Al experimentar asco, el labio superior es elevado, la nariz se arruga, mientras que los párpados inferiores se elevan y la ceja se baja. (Ekman, 2003)

### 3. Cómo funciona el lenguaje corporal

El lenguaje corporal de las personas revela que lo que dicen es comúnmente diferente a lo que piensan o sienten. Es un hecho científico que los gestos corporales revelan las intenciones de la persona. La habilidad para leer las actitudes y pensamientos a través de su comportamiento era el sistema de comunicación usado por los seres humanos previo a la existencia de los idiomas.

La gran mayoría de personas cree que el habla es la forma principal de comunicación, pero el habla ha estado presente en nuestras vidas solo recientemente, evolutivamente hablando, y nos sirve más que todo para comunicar datos y hechos. El habla se desarrolló entre medio millón y 2 millones de años atrás. Antes de este período, sonidos vocales emitidos en la garganta y lenguaje corporal eran las principales formas de comunicación y todavía lo son actualmente. La mayoría de científicos están actualmente de acuerdo en que las palabras sirven para transmisión de información, mientras que el lenguaje corporal es utilizado para negociar



actitudes interpersonales, y en algunos casos se vuelve sustituto para los mensajes verbales. El lenguaje corporal es una reflexión externa de la condición emocional de la persona. Cada gesto o movimiento que realiza la persona puede ser un indicador del sentimiento que la persona experimenta. (Pease, 2017)

Albert Mehrabian fue un investigador, pionero de la psicología del lenguaje corporal. En el año 1950, descubrió que el impacto de un mensaje transmitido de un humano a otro, es 7% verbal, 38% vocal (tonalidad inflexión, ritmo de la voz) y 55% no verbal. Ray Birdwhistell fue pionero de la comunicación no verbal. En su estudio realizó estimaciones similares a Mehrabian, concluyendo que la parte verbal en una conversación cara a cara representa menos de 35%, mientras que la parte no verbal aproximadamente 65%. (Pease, 2017)

En un análisis hecho por el autor del libro "El libro definitivo sobre lenguaje corporal" descubrieron que, según el análisis de miles de entrevistas de ventas y negociaciones grabadas durante 10 años, el lenguaje corporal tenía entre 60% y 80% de impacto en la negociación. Cuando se trataba de una negociación por llamada telefónica, la persona con mayor convencimiento sobre el argumento que presentaba, comúnmente ganaba. Pero esto no es del todo así cuando es una negociación cara a cara, puesto que la decisión final es tomada más por lo que se ve (gestos corporales) que por lo que se escucha. (Pease, 2017)

El término percepción o intuición se refiere a la habilidad de leer el lenguaje corporal de una persona y compararlo con las señales verbales. Ser una "persona intuitiva" se refiere a la habilidad de detectar contradicciones entre las palabras y el lenguaje corporal.

Se han realizado varios estudios para detectar y descubrir si las señales corporales son aprendidas, transmitidas o se nace con ellas. La evidencia en estos estudios fue recolectada de niños recién nacidos de diferentes culturas. Las conclusiones indican que cada cultura estudiada utiliza los mismos gestos faciales al nacer, por lo que se concluyó que los gestos se tienen al nacer de manera natural sin aprenderlos. Por lo tanto, existen diferencias culturales que puedan llegar a cambiar ligeramente algunos gestos, pero los gestos corporales básicos se mantienen en todas las culturas. Se debe hacer notar que la mayoría de señales comunicativas son las mismas en todo el mundo. Es decir, sonreír significa estar feliz; fruncir el ceño puede indicar estar enojado o triste. Realizar un cabeceo de arriba hacia abajo es universalmente expresarse para "decir que sí". Este es un gesto que traen los recién nacidos, puesto

que las personas que nacen sin visión también lo realizan sin haberlo aprendido. El gesto de mover la cabeza de izquierda a derecha denota "no" o negación, y es aprendido durante las etapas iniciales de la infancia. (Pease, 2017)

Algunos gestos pueden variar de cultura en cultura, así como cambia también el lenguaje. Un gesto claro de interpretar y comprender en una cultura puede ser diferente en otra.

#### 4. Gestos universales

El encogimiento de hombros, es un ejemplo universal que da a entender a la otra persona que no se comprende o no se sabe lo que se le está comunicando. Es un gesto múltiple que involucra los hombros, las manos y en ocasiones las cejas. Cuando se expresa este gesto la persona extiende sus brazos con las palmas hacia adelante y un pequeño levantamiento de cejas. Las palmas expuestas son para indicar que no se está escondiendo algo; los hombros ligeramente levantados son para proteger la garganta de un ataque y las cejas levantadas son un saludo de tipo sumiso.

#### 5. Cómo leer y analizar gestos corporales

Una regla básica es la siguiente: lo que se ve y escucha en una situación no necesariamente revela las actitudes reales que tienen las personas. Existen ciertas reglas para la lectura eficaz de gestos, estas se detallan a continuación:

a. Lectura de gestos en grupos:

Se debe analizar los gestos en conjunto con otros, tomando en cuenta el contexto de la situación del individuo. Un gesto por sí solo puede dar a entender una actitud de la persona, pero conjugado con otro o en un ambiente diferente cambia por completo de significado.

b. Búsqueda de congruencia:

Cuando las palabras y los gestos están en conflicto, es común que los humanos ignoren las palabras y se enfoquen en los gestos. Esto se debe a que las señales no verbales tienen 5 veces más impacto que la expresión verbal. Ejemplo: Sigmund Freud evaluó a una pareja de esposos. Al explicarle lo bien que iba su matrimonio, la mujer al mismo tiempo subía su anillo de casamiento de arriba a abajo. Según Freud, esto mostraba que inconscientemente las palabras no coincidían con las emociones

expresadas.

c. Lectura de gestos en contexto:

Así como el gesto de cruzar los brazos de manera apretada contra el pecho puede significar estar en una posición defensiva, bajo un contexto diferente en donde la persona se encuentra en un día de invierno, simplemente significa que tiene frío. A su vez si esa misma persona está en un ambiente en donde se le ofrece algún producto o idea del cual no está de acuerdo, estos mismos gestos pueden ser correctamente interpretados como negación o rechazo.

Es importante mencionar la dificultad de "fingir" los gestos. Si bien una persona puede alterar un gesto específico a propósito, es muy difícil expresarse congruentemente si se finge y los seres humanos (las mujeres especialmente) detectan fácilmente cualquier incongruencia. Cuando una persona quiere ser franca y sincera comúnmente expondrá las palmas de sus manos a la otra persona. Dicho gesto es completamente subconsciente y nos da una idea intuitiva de que la persona está diciendo la verdad. Al contrario, cuando por ejemplo un niño quiere esconder algo, tenderá a llevar sus manos hacia atrás de su espalda. En el caso de un adulto, lo más probable es que si quiere no revelar alguna información, lo que hará es llevar sus manos a sus bolsillos o esconderlas con un cruce de brazos. Las palmas escondidas pueden llevarnos a pensar que el individuo no está siendo honesto. En el caso de las mujeres, esto no sucede tan abiertamente, sino que en vez de mostrarlo físicamente tenderán a cambiar el tema o fingir estar ocupadas con otra actividad.

Es difícil mentir mientras se tiene las palmas de la mano expuestas, según la ley de causa y efecto. Si una persona tiene las manos expuestas está siendo honesta, el simple hecho de tenerlas abiertas hace difícil poder mentir. Esto se debe a que las emociones están directamente enlazadas con los gestos que realizamos. Otro ejemplo es una persona que al cruzar los brazos, empieza a sentirse defensiva, aunque no lo estaba previamente, gracias a la fuerte conexión entre gestos y emoción en las personas.

## 6. El poder de la sonrisa

La sonrisa y la risa son señales universales para mostrar que se está feliz. La risa comienza a los cinco meses de nacidos y la risa comienza entre el mes 5 y 6 de vida. La sonrisa es usada como un mecanismo de sumisión, como muestra hacia la otra persona que no somos una amenaza y funciona como medio para solicitar que seamos aceptados.

## 7. El uso de los brazos

Cuando un individuo cruza uno o ambos brazos a través de su pecho, una barrera se forma como un mecanismo subconsciente para bloquear una situación negativa o amenaza. Al cruzar los brazos sobre el pecho, estos se ajustan de tal forma que se queda en una posición natural. Es posible, por esta razón, que cruzar los brazos a través del pecho sea un mecanismo que traemos incorporado al nacer. Así mismo, es seguro es que cuando alguien cruza sus brazos sobre su pecho, eso se traduce a una actitud defensiva, negativa o nerviosa, o en todo caso que se siente amenazado. Es también una forma de poner una barrera entre la persona y alguien más y una forma de transmitir incertidumbre e inseguridad. La emoción básica a la que puede asociarse este gesto es al miedo. A pesar de esto y como se ha mencionado previamente, el contexto de una persona con este gesto y con cualquier gesto debe tomarse en cuenta. (Pease, 2017)

Existen variaciones sobre este mecanismo de expresión. El cruzado de brazos puede ser de un solo brazo, cruzado de brazos en donde únicamente las manos están tocándose, o las manos tocan un objeto del otro brazo como un reloj, el puño de la camisa, un brazalete en la muñeca, etc. También puede involucrar el sostenimiento de una bebida o un café o el uso de otro objeto. Aunque "la barrera" creada sea sutil en este caso, de igual manera se consigue crear la sensación de seguridad al realizar este gesto. Cualquier gesto en el que se crucen los brazos, aunque sea de manera sutil, en el contexto correcto implica nerviosismo, inseguridad y duda. Cabe mencionar que para las mujeres resulta más fácil enmascarar y esconder estos sentimientos, debido a la posibilidad de utilizar bolsas de mano, carteras o ramos de flores para justificar el "cruce de brazos."(Pease, 2017).

Otro gesto a analizar es el "abrazo a uno mismo". Este tiene la peculiaridad que en

hombres y mujeres tiende a ser diferente. El origen de este gesto está relacionado con la protección que los padres nos daban de pequeños, por medio de un abrazo o protección con su cuerpo cuando nos sentíamos inseguros o indefensos. Para sustituir esta protección en la edad adulta, las mujeres comúnmente realizan un cruce de un brazo hacia el otro, de manera que parece que un brazo está deteniendo el otro. En los hombres, los dos brazos se unen delante del cuerpo protegiendo con las manos sus partes reproductivas. El gesto está relacionado con emociones como sentirse indefenso, desprotección y vulnerabilidad.

El caso contrario a los cruces de brazo que se mencionaron, es colocar los brazos de manera en que el torso está desprotegido. Si se está sentado en una silla con una mesa, los brazos quedan en posición recta mostrando completamente la parte del pecho y estómago. Estar en esta posición está relacionado con estar abierto a la situación en que la persona se encuentra, o estar abierto a las ideas que se le están comunicando a la persona. También comunica carácter fuerte y seguridad.

Las manos revelan actitudes y emociones debido a que siempre están por delante de nosotros. Cuando hablamos de las manos, existe un gesto que todos hemos visto o usado en nuestras vidas. Es el gesto de las manos apretadas entre sí. Este gesto, a pesar de que puede parecer que transmite confianza y seguridad, es una muestra de actitud negativa, ansiosa o una muestra de frustración, incluso si la persona sonrío. En negocios se ha examinado dicha postura, y detectado que se tiene a implementar cuando el individuo siente que no está siendo convincente o que está perdiendo la negociación. Hay tres tipos de este gesto: el primero es con las manos atrincheradas enfrente de la cara, el segundo es el de las manos atrincheradas sobre un escritorio, y el último sobre la entrepierna cuando se está parado. (Pease, 2017)

El uso de las manos enfrente del cuerpo, realizando una frotación de arriba hacia abajo puede significar expectativa de tipo positivo. Como alguien que espera una buena noticia, experiencia o un regalo. También está el gesto del "campanario", que se basa en el posicionamiento de los dedos de una mano sobre los dedos de la otra, formando una especie de "campanario". El gesto puede incluir el movimiento de una mano sobre la otra, de tal forma que se observa como si una mano "le hace presión" a la otra. Este gesto es comúnmente utilizado en una situación en donde la persona está en superioridad e indica una actitud de confianza y seguridad en sí mismo.

Tener las manos sostenidas una a la otra por detrás de la espalda es un gesto que hemos visto antes en policías, guardias, directores de colegio, comandantes y en

general cualquier persona en una posición de poder. La persona al estar en esta posición, expone su pecho, estómago, corazón, entrepierna, garganta y cabeza, quedando vulnerable y comunicando implícitamente una actitud de no tener miedo. Este gesto denota emociones de superioridad, seguridad, confianza y poder. Una variación de esta posición es cuando una palma procede a sostener al brazo opuesto, específicamente en la parte del antebrazo. En esta situación, la persona no comunica seguridad sino sentimientos de nerviosismo. (Pease, 2017)

## 8. El uso de los pulgares

Cualquier posición en donde se resguardan las manos y los pulgares quedan visibles, puede implicar actitudes de superioridad, confianza, tranquilidad, dominancia de una situación. Cuando las manos quedan inmersas en los bolsillos del pantalón o chaqueta, el diferenciador para evaluar si la persona se siente confiada es la muestra de los pulgares. Ya sea que las manos estén delante o detrás en los bolsillos del pantalón, la muestra de los pulgares es una señal positiva, utilizada por un individuo que se siente a gusto. Los gestos, como se ha mencionado, deben leerse en grupos. Para el caso del ejemplo anterior, si la persona cruza los brazos manteniendo los pulgares visibles, el análisis debe hacerse en conjunto. Este gesto significa que la persona está cerrada pero aun así se siente en posición de superioridad.

## 9. Gestos de procrastinación

Cuando una persona comienza a llevar sus manos a la cabeza es una señal de que está comenzando a aburrirse y utiliza una mano para detener su cabeza y así evitar quedarse dormido. Que tan aburrido esté una persona está relacionado con que tanto utiliza su mano para detener su cabeza. Es decir, si está completamente apoyando la cabeza sobre la mano, el nivel de aburrimiento es alto. Por otra parte, realizar toques continuos sobre la mesa con los dedos o los pies, son comúnmente interpretados por la gente como aburrimiento, pero más bien se trata del sentimiento de impaciencia. El nivel de aburrimiento en este caso se ve reflejado en que tan rápido la persona mueve sus pies o realiza toques a la mesa.

## 10. Gestos de evaluación

Cuando se está evaluando una situación, se tiende a colocar la mano sobre la mejía, usualmente con el dedo índice apuntando hacia arriba. Si se está interesado se tenderá a mantener la cabeza recta y no apoyada en la mano. Cuando del estado de evaluación se pasa a aburrimiento, se comienza a asentar la cabeza en la palma de la mano, con el dedo aun apuntando hacia arriba. Este mecanismo se utiliza con el objetivo de aparentar estar interesado cuando en realidad no se está. Si el dedo está en posición apuntando hacia arriba, pero la cabeza empieza a asentarse sobre la mano, esto se asocia con actitudes o pensamientos negativos sobre lo que se está escuchando. (Pease, 2017).

Los gestos de llevar las manos al frente de la cabeza o a la parte de atrás del cuello están asociadas con miedo, angustia o frustración. Se ha demostrado que las personas que tienden a llevar sus manos a la frente cuando experimentan estas emociones mencionadas, son más abiertas y llevaderas que las personas que las llevan a la parte anterior del cuello, que más probablemente tienen una tendencia a la negatividad y a ser más críticas.

## 11. La importancia de las piernas en el lenguaje corporal

Se ha estudiado sobre el lenguaje corporal que mientras más alejada está del cerebro una parte del cuerpo, menos atención se le da para controlar lo que dicha parte está haciendo. Como la cara está muy cerca, tenemos comúnmente control de los gestos que realizamos e incluso un repertorio de gestos fingidos para situaciones determinadas. En el caso de las piernas ocurre lo contrario, puesto que al estar alejadas muchas veces no se le da atención de cómo están interactuando en nuestro cuerpo. El cerebro está programado naturalmente para gravitar hacia todo aquello que queremos y evitar todo lo que no. Es por eso que las piernas y pies revelan lo que se quiere hacer y a dónde se quiere ir. Por ejemplo, cuando se tienen las piernas abiertas y sin cruzar indican una actitud dominante y segura mientras que cuando las piernas se cruzan revelan que la persona está cerrada o insegura. Existen varias posiciones de las piernas. A continuación, se verán sus respectivos significados y las emociones que involucran:

### a. La posición de "atención":

Esta posición es neutral e indica que la persona no está pensando ni en quedarse

ni en irse, simplemente neutral. Es utilizada por niños que le hablan al profesor, empleados que le hablan a su jefe y soldados que le hablan a su comandante. En general, usada cuando la persona establece una comunicación con una persona de status más alto en un ambiente determinado.

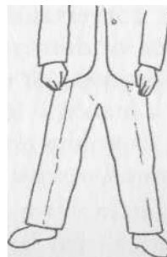
Figura 1: Posición de atención



b. La posición "piernas abiertas":

Esta posición es usada principalmente por hombres e indica claramente que el individuo no está dispuesto a irse del lugar. En esta posición los pies están plantados y comunican autoridad y dominancia, puesto que en hombres resalta la parte de los genitales.

Figura 2: Posición piernas abiertas



c. La posición "Un pie adelante":

Esta posición se caracteriza por tener un pie puesto firmemente en el suelo, dejando caer todo el peso sobre este pie. La postura demuestra la intención de la persona, a través de la posición del pie que no está firme en el suelo. Si la persona inclina dicho pie hacia una posición específica, ahí es a donde tiene intenciones de ir.



Figura 3: Posición pie adelante



d. La posición "Piernas cruzadas":

Es una posición muy característica de personas que están en un ambiente en donde no conocen demasiado a las otras personas, es decir no están familiarizadas con la gente del ambiente en que se encuentran. Al contrario de la posición de piernas abiertas que demuestra apertura, esta demuestra que el individuo está cerrado y tiene una actitud defensiva puesto que los genitales están simbólicamente cubiertos por las piernas.

Figura 4: Posición piernas cruzadas



Si ambas piernas y brazos están cerrados, es un indicador que dicha persona no está dispuesta a hablar o interactuar, teniendo una actitud cerrada y defensiva.

## 12. Posiciones básicas de la cabeza

a. Posición "Cabeza arriba":

Cuando la cabeza está en posición neutral, o inclinada ligeramente hacia arriba, la persona tiene una actitud neutral ante la situación que experimenta o lo que está escuchando. Cuando la posición cambia y la persona está con la cabeza como si estuviera viendo hacia arriba, la posición está relacionada con actitudes de superioridad, audacia o incluso arrogancia. El hecho de posicionar la cabeza de tal forma que la garganta está visible hacia la otra persona, está asociado con la agresividad y poder.

Figura 5: Posición cabeza arriba



b. Posición "cabeza hacia un lado":

Cuando el individuo tiene esta posición, la garganta y el cuello están expuestos y visibles, viéndose hacia los demás como más pequeño y menos autoritario. El origen de esta posición puede provenir de la forma en que un recién nacido recuesta su cabeza sobre su madre. Estar en esta posición implica mostrar debilidad, sumisión y vulnerabilidad. Es en su mayoría utilizada por mujeres, como muestra de sumisión.

Figura 6: Posición cabeza hacia un lado



c. Posición "cabeza abajo":

Estar en esta posición, en la cual la barbilla apunta hacia abajo y la garganta no se muestra, implica actitudes relacionadas con negatividad, actitud defensiva o agresiva y crítica negativa. También muestra desaprobación y abatimiento. Por ejemplo, una evaluación crítica realizada por una persona sobre un acontecimiento, es comúnmente realizada en esta posición.

Figura 7: Posición cabeza abajo



d. Posición encogimiento de hombros":

Esta posición se da cuando se elevan los hombros ligeramente y la cabeza se baja, con el objetivo de proteger el cuello y garganta. En un contexto social implica que la persona siente inferioridad y puede ser generada en la persona cuando esta quiere disculparse de forma sumisa. Es utilizado en el ambiente laboral cuando un empleado se siente inferior a sus superiores.

Figura 8: Posición hombros encogidos



e. Posición "manos sobre la cintura":

Es el gesto universal que comunica que se está preparado para la situación y para tomar acción. En ciertas culturas puede tener una connotación de agresión y enojo. No es necesario tener las dos manos para comunicar que se está listo para la acción, una mano en la cintura envía a los que ven el gesto, la misma señal. Comúnmente los hombres utilizan la postura para demostrar una actitud asertiva masculina alrededor de una mujer. Esto tiene sentido dado a que cuando entramos en esta posición, naturalmente se abarca más espacio en el ambiente, lo cual da la apariencia que se es más grande.

Figura 9: Posición manos sobre la cintura



Existe una versión similar cuando la persona está sentada, y en esta versión las manos van en la parte anterior de la cabeza, y transmite que el individuo está relajado o confiado, aunque también podría transmitir que la persona experimenta una falsa apariencia de seguridad.

f. Postura del vaquero:

Consiste en que los dos dedos pulgares están inmersos en el pantalón, cincho o bolsillos y los demás dedos de la mano quedan sostenidos en el pantalón. La posición encuadra los genitales y es utilizada mayoritariamente por hombres para mostrar una actitud sexual agresiva. Utilizada para demostrar valentía entre hombres, y actitud de dominancia. Cuando una mujer utiliza la postura comunica su asertividad sexual. (Pease, 2017)

Figura 10: “Postura del vaquero”



## B. Aprendizaje de máquina

“Machine learning” o aprendizaje de máquina es un área que involucra varios procesos, que a su vez hacen difícil de definir el término. “Aprender” para una máquina puede tener diversos significados. Generalmente se le llama aprendizaje de máquina al hecho que una máquina altere un programa o sus datos, dependiendo de las entradas y estímulos que se le provee. (Russell, S. J.2003)

Al mencionar la definición, se plantea la pregunta de por qué es necesario que la máquina aprenda y no que el programador implementa el programa para que la máquina funcione como se espera, desde el principio. Como primer punto, y mencionando el aspecto humano, el hecho que una máquina aprenda nos puede proveer diversas percepciones de cómo se da el aprendizaje en humanos y animales, y una mejor idea de los procesos utilizados para el aprendizaje.

Por otra parte, y mencionando la parte tecnológica ahora, existen varios motivos por los cuales es preferible utilizar “aprendizaje de máquina”. Existen varias tareas que los humanos realizamos que, aunque sean automáticas o sencillas de hacer para nosotros, no son nada fáciles de replicar para una máquina. Por ejemplo, manejar un carro. Esta es una tarea que involucra múltiples procesos y muchísimas decisiones tomadas en pocos segundos. Decirle a un programa de manera explícita todas las posibles variables y situaciones en este caso, sería demasiado complejo y prácticamente imposible.

Existen tareas o procedimientos que no pueden ser definidos de manera correcta. Es decir, se tiene conocimiento de las entradas y salidas del programa, pero no hay una relación bien definida entre las entradas y salidas que se esperan; Para esto es útil que las máquinas utilicen algoritmos de aprendizaje, para proveer salidas correctas en base a la relación detectada entre los datos. También nos ayuda el “aprendizaje de máquina” cuando nos enfrentamos a tareas que no están al alcance de las capacidades humanas, como cuando se quiere analizar grandes cantidades de datos (big data). Por ejemplo, el análisis de datos de la población de los Estados Unidos con el fin de encontrar las enfermedades más comunes según edad y etnicidad. (Russell, S. J.2003).

Los algoritmos de “machine learning” son un campo de estudio de la inteligencia artificial, que se basa en la creación de modelos predictivos que se generan al analizar grandes cantidades de datos. El objetivo de este campo es lograr construir un agente inteligente que pueda resolver problemas sin necesidad de programarlo de manera explícita. La máquina aprende a través de un conjunto de reglas abstractas, que se auto configuran a partir de las entradas y forma de los datos. (Russell, S. J.2003)

En el área de “aprendizaje de máquina”, existen varios tipos de algoritmos, entre los que se encuentran los algoritmos de clasificación y algoritmos de regresión. Los primeros son utilizados cuando el resultado a obtener es de tipo discreto, debido a que proveen como resultado una clase a la que pertenece un dato. Los últimos se utilizan para predecir resultados de tipo continuo. El algoritmo de “aprendizaje de máquina” utilizado en este trabajo, se llama “Algoritmo de regresión logística, que es un algoritmo básico de “aprendizaje de máquina” y análisis estadístico, siendo uno de los más utilizados en análisis de datos.

## 1. Algoritmo de regresión logística

Se utiliza para problemas de clasificación binaria (dos clases de valores). Este método también es llamado modelo logístico o “logit model”. El algoritmo analiza la relación entre múltiples variables independientes y una categórica variable dependiente, y estima la probabilidad de ocurrencia de un evento mediante el ajuste hacia la curva logística. Existen dos tipos de regresión, de tipo binaria y de tipo multinomial. La regresión logística binaria es utilizada cuando la variable dependiente es dicótoma y las variables independientes son continuas o categóricas. Cuando las variables dependientes no son dicótomas y existen más de dos categorías o clases, se puede utilizar la regresión logística multinomial. (Park, 2013)

La regresión logística se basa en la función logística, de ahí proviene el nombre del algoritmo. Esta función se conoce también como la función sigmoide, que asegura que un resultado específico esté siempre entre 0 y 1.

Figura 11: Función sigmoide

$$y = \frac{1}{1 + e^{-z}}$$

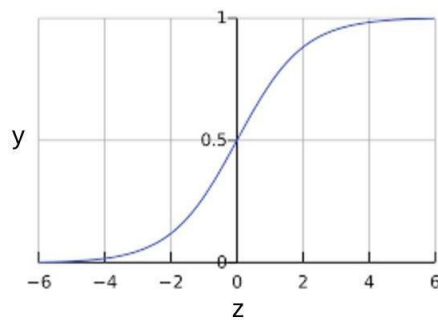
La ecuación sigmoide que se puede visualizar arriba, el valor de  $y$  representa una probabilidad, mientras que  $z$  representa el resultado de la capa lineal de un modelo entrenado con regresión logística. Para representar la información, la regresión logística utiliza una ecuación. En la ecuación, los valores de  $x$  son las entradas las cuales se combinan con pesos, que representan una probabilidad de que tan certero es el valor, con lo cual se obtienen las salidas con valor  $y$ . La ecuación se ve a continuación, en la cual del lado izquierdo  $P$  representa la probabilidad.

Figura 12: Ecuación de regresión logística

$$\ln\left(\frac{P}{1-P}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k$$

La función sigmoide se puede visualizar en la Figura 11. Claramente se observa que los valores de  $y$  adoptan como resultado un valor entre 0 y 1.

Figura 13: Gráfica de la función sigmoide



## 2. Supuestos del algoritmo de regresión logística

La primera suposición del modelo de regresión logística es que la regresión logística no asume errores en la variable de salida  $y$ . Esto hace que los datos a ingresar al modelo deben estar clasificados de manera óptima y evitar datos aislados. Un concepto a entender previo a analizar las suposiciones, es el “maximum likelihood”. Este es un método que sirve para estimar los parámetros de un modelo. El método selecciona un conjunto de valores de los parámetros del modelo que maximizan la función de “likelihood”. (Hurlin, 2013)

Otras suposiciones que aplican al modelo de regresión logística son:

Se requiere que la variable dependiente sea discreta, y mayoritariamente dicótoma; Debido a que la regresión logística estima la probabilidad de un evento ocurriendo  $P(Y=1)$ , es necesario codificar la variable dependiente como corresponde. Es decir, que la salida deseada debe ser codificada como 1.

El modelo debe estar ajustado correctamente, no puede estar sobre ajustado con las variables insignificantes incluidas. Tampoco debe estar sub ajustada mediante la exclusión de variables significativas.

El modelo no debe tener multicolinealidad (se logra a través de no relacionar ninguna variable predictora con otras variables) y requiere que cada observación sea independiente, es decir, que las variables independientes no sean funciones lineales de las otras.

Como último punto, se requiere que las muestras de datos sean bastante grandes, debido a que las estimaciones del maximum likelihood son menos potentes que los cuadrados utilizados en la regresión lineal. (Park, 2013)

### 3. Estimación humana de poses

Una pose se puede definir como la configuración o estructura que adopta el cuerpo humano en un instante determinado. Un ejemplo de una pose es la configuración de las articulaciones de una persona cuando tiene las manos en la cabeza. La estimación humana de poses se refiere a las técnicas de visión computacional, que detectan figuras humanas a partir de una imagen o video, con el objetivo de determinar las partes del cuerpo. El algoritmo de esta técnica, detecta las principales articulaciones del cuerpo, pero no información de quién es la persona u otra información asociada a la persona.

A partir de esta estimación, se pueden desarrollar aplicaciones útiles para el ser humano, involucrando realidad virtual, aplicaciones deportivas y animación de películas.

En la actualidad, se han desarrollado varios algoritmos para detección de poses. En años anteriores, para utilizar software de estimación de poses se necesitaba equipo de hardware complejo. Actualmente, esto ha cambiado y basta con tener un ordenador con una webcam y un navegador instalado en el ordenador. Además, de las librerías desarrolladas sobre estimación de poses en los últimos cinco años, existen varias que son de código abierto, para utilización y desarrollo de cualquier persona.

La estimación de poses sucede en dos fases principales: la primera es la alimentación de una imagen hacia una red neuronal convolucional. La segunda parte es la utilización de un algoritmo de decodificación de poses simples o múltiples con lo que se obtienen los puntos claves “keypoints” de las articulaciones, las ponderaciones de cada par de puntos, y el valor de confianza en general de la imagen.

En la sección siguiente, Algoritmo Pose Net, se puede encontrar la descripción en detalle de las entradas que se le provee a un algoritmo de estimación de poses humanas. (Assgoba, 2018). Entre otras librerías que se pueden utilizar para la estimación de poses están: Dense Pose, que mapea píxeles del cuerpo. (Neverova, 2018). Y TF-Open-Pose que detecta puntos clave de la cara, pies y manos.



Figura 14: Puntos clave de una estimación de pose.



#### 4. Algoritmo PoseNet

El algoritmo PoseNet es un algoritmo desarrollado en el año 2018, por científicos de Google. (Oved, 2018) El software está implementado en TensorFlow, una librería de JavaScript para entrenamiento e implementación de modelos de “aprendizaje de máquina” en el navegador. La herramienta PoseNet recibe una imagen o video a través del navegador y permite la estimación de las poses del cuerpo humano en tiempo real. El resultado del algoritmo es una imagen con un esqueleto sobrepuesto que modela las articulaciones del cuerpo humano, y un conjunto de puntos etiquetados con los nombres de las principales articulaciones.

El algoritmo permite estimar dos tipos de poses, simples y múltiples. Las poses simples son aquellas en donde solo una persona es analizada por el algoritmo. En cambio, en las poses múltiples, el algoritmo analiza más de una persona, o grupo de personas. Como el enfoque de este trabajo está en analizar a una persona individual, se utiliza el algoritmo de poses simples.

El resultado de los puntos del cuerpo etiquetados se llaman punto clave “keypoints”. Un punto clave puede ser la “oreja derecha” o la “muñeca izquierda”. Existen 17 puntos claves en total y todos tienen una identificación que comienza de cero y llega hasta 16, además del nombre de la parte del cuerpo en texto plano. A continuación se

observa esta tabla de puntos clave, solo se muestra hasta el id número 10 por propósitos de visualización:

Figura 15: Tabla de puntos clave

<b>Identificador</b>	<b>Parte</b>
0	nariz
1	ojoIzquierdo
2	ojoDerecho
3	orejaIzquierda
4	orejaDerecha
5	hombroIzquierdo
6	hombroDerecho
7	codoIzquierdo
8	codoDerecho
9	muñecaIzquierda
10	muñecaDerecha
11	caderaIzquierda
12	caderaDerecha
13	rodillaIzquierda
14	rodillaDerecha
15	tobilloIzquierdo
16	tobilloDerecho

## 5. Funcionamiento del algoritmo PoseNet

A través del análisis de diferentes algoritmos, se detectó que este es un algoritmo ideal para la resolución del problema de este trabajo. El algoritmo, como se mencionó previamente permite la obtención de un conjunto de puntos en el plano x,y junto con un puntaje de confianza (score) que indica la probabilidad de que el algoritmo detectó la parte del cuerpo en esa posición x,y de forma correcta. Como entradas el algoritmo recibe una imagen o conjunto de imágenes, un valor entre 0.2 y 1, en donde 1 indica que la imagen quedará al 100% de su tamaño original y 0.2 al 20% (para reducir el tamaño original de la imagen), un valor de “separación de salida”, el cual puede ser 8, 16 o 32, y que determina la precisión del algoritmo (a mayor precisión, menor rapidez de ejecución) y un valor booleano para la posición horizontal o vertical de la imagen. La salida que provee el algoritmo es un arreglo de poses, con los puntajes de confianza (score) y los puntos clave de cada parte del cuerpo de cada pose. La figura 16 muestra una salida parcial del algoritmo.

Figura 16: resultado del análisis de imagen

```
{
  "score": 0.32371445304906,
  "keypoints": [
    {
      "position": {
        "y": 76.291801452637,
        "x": 253.36747741699
      },
      "part": "nose",
      "score": 0.99539834260941
    },
  ],
}
```

## V. MARCO METODOLÓGICO

La primera parte del trabajo se basó en la recopilación de las imágenes de las personas en un ambiente determinado, y a través de plataformas en línea. Como primer ambiente se procedió a trabajar en la recolección de imágenes de personas, tomando fotografías de personas expresando diferentes emociones. También se recopilaron imágenes que estén disponibles libremente en plataformas digitales.

La segunda parte del desarrollo del proyecto se basó en la definición del modelo humano que se utilizó. Esto se realizó a través de un modelo de esqueleto kinésico, el cual es una representación del cuerpo humano, que consta de puntos representando las articulaciones, en donde ocurre el movimiento de las extremidades.

La tercera parte involucró el “tagging” o etiquetado de las imágenes recopiladas previamente. Se utilizó información de los gestos faciales de una persona, como base teórica para detectar la emoción de una persona. El motivo es porque los gestos faciales representan una gran parte de la expresión en el ser humano. A partir de esta idea, se buscó que el algoritmo implementado aprendiera sobre la emoción completa en el cuerpo tomando como base la emoción que se detecta en las expresiones faciales.

Se le colocó una etiqueta a cada una de las imágenes, de la emoción o conjunto de emociones que más se asociaba a lo que la persona mostraba en la imagen según sus gestos corporales y faciales. Se construyó un compilado o base de datos de las imágenes etiquetadas. Esta base de imágenes es sumamente importante, puesto que se utilizó como entrada para cualquiera de los algoritmos de “aprendizaje de máquina” implementados funcionaran correctamente. Cabe mencionar que se procedió a validar las imágenes etiquetadas con la ayuda de la licenciada en psicología Lorena Gómez. Esta validación consistió en comparar las etiquetas de 15 imágenes contra las etiquetas que la licenciada consideró para cada imagen. (La emoción predominante en cada imagen)

Después de haber realizado esto, se procedió a implementar la solución de software, utilizando teoría de “aprendizaje de máquina” e inteligencia artificial. Se utilizaron los datos recopilados para el desarrollo de la inteligencia artificial del proyecto.

Entre los recursos físicos utilizados, se encuentran los siguientes: una computadora portátil hp Intel Core i5 con 8 GB de memoria RAM, para el desarrollo de software con capacidad para análisis y procesamiento de imágenes, una cámara de un iPhone 7 con 12 megapíxeles para captar personas claramente. Entre los recursos virtuales, se

utilizó Adobe Photoshop para la edición de imágenes. Se considera que no es necesario ningún otro recurso físico o virtual.

## A. Recopilación de imágenes y etiquetado

Para poder desarrollar este proyecto, se compiló un conjunto de imágenes de personas, en las que claramente se observa una emoción por cada imagen. Este compilado de imágenes es nuestra “base de datos”, que nos sirve para alimentar al algoritmo de “aprendizaje de máquina” utilizado. Todas las imágenes recolectadas, están libres para uso personal, y no tienen restricciones de derechos de autor.

Las características de las imágenes recolectadas fueron las siguientes: postura frontal de la persona con las partes del cuerpo claramente visibles; la profundidad de las imágenes no fue tomada en cuenta para el análisis, para simplificación del problema. Además de esto, se utilizaron únicamente imágenes que tuvieran licencia de creative commons o libres para uso personal. Y, por último, se analizaron solamente poses estáticas, no se incluyeron poses en movimiento.

Las imágenes recolectadas se seleccionaron en base a la teoría de emociones presentada en previamente. A partir de la utilización de dicha información, se seleccionaron fotografías de personas que representaran una de las seis emociones básicas (felicidad, tristeza, enojo, disgusto, miedo y sorpresa). Posterior a recolectar las imágenes, se procedió a etiquetar las mismas, con una de las siguientes seis etiquetas: “felicidad”, “miedo”, “enojo”, “asco”, “tristeza”, y “sorpresa”. En total se etiquetaron 400 imágenes, teniendo aproximadamente 17% de imágenes etiquetadas de cada emoción. Para obtener el compilado de imágenes, se utilizaron las siguientes seis imágenes que se pueden ver a continuación, como base para encontrar las imágenes adecuadas para agregar al compilado. Estas se observan en las seis figuras siguientes:

Figura 17: Emoción de sorpresa



Figura 18: Emoción de miedo



Figura 19: Emoción de enojo

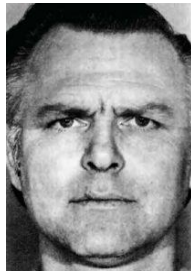


Figura 20: Emoción de felicidad



Figura 21: Emoción de tristeza



Figura 22: Emoción de disgusto



El problema a resolver involucra predicción, a partir de un conjunto de datos discretos. Para obtener los resultados, las emociones fueron mapeadas hacia los siguientes valores: felicidad: 0, miedo: 1, tristeza: 2, enojo: 3, sorpresa: 4, disgusto: 5

Es importante hacer notar que el etiquetado fue realizado conforme a la emoción visible en la cara de la persona, puesto que se desea que el algoritmo de aprendizaje de máquina aprenda sobre el lenguaje corporal (sin incluir los gestos de la cara).

#### A. Elección del modelo de clasificación de emociones

Debido a que se desea predecir una emoción específica y única en cada una de las imágenes de nuestro compilado de imágenes, el tipo de modelo a utilizar es de tipo categórico (en donde cada una de las posibles emociones representa una categoría). Esto se fundamenta en que se pueden utilizar las etiquetas de las emociones de las imágenes como una clase o categoría.

Las emociones fueron seleccionadas en base al modelo de Paul Ekman definido en libro “*Unmasking the face*”, que indica que las emociones básicas son: felicidad, tristeza, enojo, disgusto, miedo y sorpresa. Se seleccionó el de Paul Ekman, puesto que, en otros modelos de emociones, existen emociones que son muy similares. Esto causaría que al clasificar las imágenes se tendría el problema de qué emoción asignarle a una imagen cuando pueden ser dos las emociones posibles. Con el modelo de Ekman, las emociones son bastante generales. Existen estudios recientes que han reducido las emociones básicas a solo cuatro, combinando cómo una única emoción el miedo y la sorpresa, además del enojo y disgusto.

### C. Obtención de los datos

Para obtener los datos de la estimación de las poses de las imágenes, se desarrolló una aplicación web sencilla utilizando Node JS y JavaScript (ver código en anexos). En esta aplicación, se utilizó la librería de PoseNet para la obtención de los resultados de cada imagen (datos de puntos clave y ponderaciones). Se limitó la obtención de los puntos clave, solo desde la cadera hacia arriba. Los puntos clave de las rodillas y pies no se consideraron, como se puede observar en la Figura 23. Esto se hizo de esta manera ya que en el texto “El libro definitivo sobre el lenguaje corporal” el autor menciona que las piernas no comunican tanta información relevante como el torso, para decidir una emoción en una persona.

El resultado de cada imagen se obtuvo de la forma como se observa en la Figura 23, y esto se hizo para cada una de las imágenes. (En la Figura 23, se observa el resultado para una imagen etiquetada como “happy”. El resultado provee las coordenadas x,y de la posición de la parte del cuerpo)

Figura 23: Salida de la aplicación con una imagen etiquetada “felicidad”

Part	Happy2	
	x	y
nose	903.4765542	275.8524654
leftEye	937.726755	244.046665
rightEye	862.3764363	254.009373
leftEar	993.0343506	271.9609889
rightEar	823.2991225	278.6326712
leftShoulder	1102.141458	521.5452937
rightShoulder	727.394206	539.2096522
leftElbow	1146.320054	839.5583757
rightElbow	633.3458538	837.44614
leftWrist	1142.046279	1110.427119
rightWrist	656.940081	1081.123177
leftHip	999.2870039	1105.634306
rightHip	767.0110112	1088.992209



## D. Normalización de los datos

Luego de obtener todos los datos, se procedió a normalizar los mismos. Esto se realiza por varios motivos: primero, las personas en las imágenes no están todas en la misma posición y no tendría sentido analizar un punto de coordenadas de una imagen contra un punto de coordenadas de otra imagen; segundo, las personas en las imágenes no están todas en la misma escala, por lo que se deben modificar la escala de las imágenes para que coincidan los puntos clave de cada persona con los de otra. En resumen, la normalización realizada involucra la traslación y cambio de escala de las imágenes.

Primero se cambiaron de escala las imágenes, para lo cual se utilizó la distancia entre los ojos como métrica. Es decir, se restó el valor x,y del ojo izquierdo con el valor x,y del ojo derecho, para obtener la distancia de un ojo a otro. Esta distancia se dividió dentro de todos los otros valores del cuerpo; y así con todas las imágenes. Para la parte de traslación, se trasladaron todas las poses de las personas a la posición 0,0 de la imagen, para lo cual simplemente se restó la distancia del ojo derecho tanto en x y y, para todas las imágenes.

Es importante mencionar, que se seleccionaron imágenes en donde las personas se encuentran en posición frontal para asegurar que los datos fueran más sencillos de normalizar y así evitar errores. Por esta razón la rotación de las personas en las imágenes no forma parte de la normalización realizada. El resultado final de la normalización se puede ver en la Figura 24.

Figura 24: Datos sin normalizar y normalizados

Datos sin normalizar			Datos Normalizados		
Parte	Feliz		Parte	Feliz	
	x	y		x	y
nariz	903.4765542	275.8524654	nariz	0.545453803	0.2898871935
ojoIzquierdo	937.726755	244.046665	ojoIzquierdo	1	-0.1322185247
ojoDerecho	862.3764363	254.009373	ojoDerecho	0	0
orejaIzquierda	993.0343506	271.9609889	orejaIzquierda	1.734006126	0.2382420697
orejaDerecha	823.2991225	278.6326712	orejaDerecha	-0.5186084737	0.32678426
hombroIzquierdo	1102.141458	521.5452937	hombroIzquierdo	3.182004084	3.550561233
hombroDerecho	727.394206	539.2096522	hombroDerecho	-1.79139561	3.784991011
codoloIzquierdo	1146.320054	839.5583757	codoloIzquierdo	3.768313425	7.771022237
codoDerecho	633.3458538	837.44614	codoDerecho	-3.039543645	7.74299003
muñealIzquierda	1142.046279	1110.427119	muñealIzquierda	3.711594691	11.3658145
muñecaDerecha	656.940081	1081.123177	muñecaDerecha	-2.726416539	10.97691182
caderalIzquierda	999.2870039	1105.634306	caderalIzquierda	1.816987238	11.30220743
caderaDerecha	767.0110112	1088.992209	caderaDerecha	-1.265627361	11.08134444

## E. Discretización de los datos

Para discretizar los datos obtenidos en los pasos anteriores, se planteó la idea de utilizar una malla o una división cuadrículada de la imagen. Es decir, cada imagen se dividió en una matriz de valores de  $14 \times 14$ , siendo el tamaño de cada punto en la matriz, el tamaño de la imagen dividido 14.

Con estos se crean sectores en la imagen, en los cuales pueden contener cada parte del cuerpo en las diferentes imágenes. Esto nos sirven para clasificar cada punto  $x,y$  que representa un punto clave, para todas las imágenes. El cuadro mencionado se puede observar en la Figura 13.

Figura 25: Malla para discretización



Luego de normalizados los datos, el siguiente paso es prepararlos para poder tenerlos en el formato adecuado necesario para alimentarlos al algoritmo de regresión logística. Debido a que el tipo de algoritmo de regresión logística para este trabajo es binomial, ya que deseamos obtener un verdadero o falso sobre si la persona experimenta una emoción, necesitamos tener los datos extraídos en pasos anteriores en dos clases o categorías.

Para encontrar la primera categoría, se ideó una manera de transformar los datos. Esta se basa en el problema de que los datos extraídos por el algoritmo, tienen tres dimensiones por cada estimación de pose. Estas tres dimensiones son las coordenadas  $x$ ,  $y$ , y la parte del cuerpo de ese punto, por lo que tenemos un arreglo de valores de la forma  $(parte, x, y)$ . Esto no es ideal para nuestros fines, puesto que se seleccionó el algoritmo de regresión logística binomial, y de tratar los datos de esta manera tendríamos cuatro categorías por cada pose. Por esta razón se procede a implementar

una técnica para convertir los datos, que se explica a continuación. Se desea encontrar una forma de convertir la coordenada x,y, además de la parte del cuerpo, a un solo valor. Para esto, se utiliza la separación por cuadrantes de las imágenes, mediante la malla visible en la Figura 13. Cada cuadrante se nombra con un valor, de A hasta K, dependiendo del incremento que se utilice. En este caso se usó un incremento de dos, desde -14 a 14. Por medio de la malla, obtenemos rangos con un incremento de dos, aplicamos dichos rangos a todos los puntos de todos los puntos clave de cada pose, con lo que se encuentra en qué cuadros de la malla está localizada cada parte según la pose. Con este método, se obtiene un resultado de la forma:

## ojoIzquierdoDG

Este resultado indica que, para una pose, el ojo izquierdo se encuentra en el cuadrante DG. Se hizo esto para todos los puntos, con lo que obtenemos un dato de la forma en que se visualiza en la Figura 26, que ya están listos para alimentarlos al algoritmo de “aprendizaje de máquina”:

Figura 26: Datos finales discretizados

Cadena	Cadena Final
narizHH	narizHH,enojo
ojoIzquierdoHG	ojoIzquierdoHG,enojo
ojoDerechoGHGH	ojoDerechoGHGH,enojo
orejalzquierdaHH	orejalzquierdaHH,enojo
orejaDerechaGH	orejaDerechaGH,enojo
hombrolzquierdoli	hombrolzquierdoli,enojo
hombroDerechoGI	hombroDerechoGI,enojo
codolzquierdoliK	codolzquierdoliK,enojo
codoDerechoFK	codoDerechoFK,enojo
muñecalzquierdaL	muñecalzquierdaL,enojo
muñecaDerechaGL	muñecaDerechaGL,enojo
caderalzquierdaL	caderalzquierdaL,enojo
caderaDerechaHL	caderaDerechaHL,enojo

Cabe mencionar que la data de las imágenes etiquetadas, se ingresó al algoritmo de aprendizaje de máquina de manera equitativa. Se ingresó como entrada la misma cantidad de información de cada emoción, obtenida de las imágenes. Como primera salida del algoritmo, se obtiene un arreglo de predicciones. Este se obtuvo mediante la obtención previa del modelo de regresión logística y el ajuste de los datos. Los datos

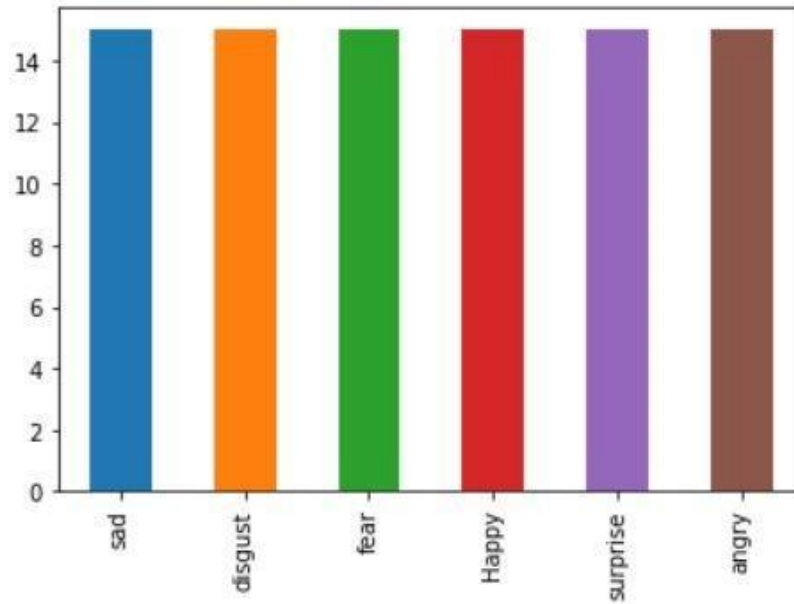
ajustados fueron el conjunto de entradas  $X$  y el conjunto de salidas  $Y$ , las cuales sabemos son la data de las partes del cuerpo en la forma “parteAB” y las emociones etiquetadas a el conjunto de partes, respectivamente.

## VI. RESULTADOS

### A. Conjunto de datos de entrenamiento

Los datos recolectados de las imágenes se ingresaron al algoritmo de regresión. La siguiente gráfica muestra el número de imágenes de cada emoción.

Figura 27: Gráfica de data clasificada (Eje X: Número de imágenes)



### B. Matriz de predicciones

Figura 28: Matriz de predicción de emociones para cada imagen

```
predictions = model.predict(first)
print(predictions)
[0 0 3 0 2 0 0 0 0 0 0 2 0 0 0 2 2 2 3 2 2 2 2 2 2 0 2 2 2 2 3 3 3 3 3 3 2
0 2 0 2 3 3 0 3 1 1 1 1 1 1 1 1 1 4 1 1 4 4 1 4 4 4 1 4 4 4 4 1 4 4 4 4 1
4]
```

### C. Puntaje/precisión del modelo

Mediante la función `score()` de la data X de entradas y Y de salidas (en este caso “first” y “second”), obtenemos el score del modelo, que nos muestra la precisión media de las predicciones.

Figura 29: Resultado del modelo

```
model.score(first,second)
0.7733333333333333
```

### D. Resultado de regresión logística

Figura 30: Regresión logística

```
Logistic Regression: 0.565657
```

### E. Matriz de confusión

Figura 31: Tabla de matriz de confusión

```
[[2 0 0 1 0]
 [0 7 0 0 0]
 [0 0 7 1 0]
 [1 0 0 1 0]
 [0 0 0 0 3]]

Confusion Matrix
```

### F. Precisión del modelo

Como se observa en la siguiente figura, el modelo nos provee un resultado de precisión, para saber el porcentaje de aciertos utilizando el conjunto de datos de validación cruzada, y las predicciones obtenidas.

Figura 32: Precisión del modelo

```
predictions = model.predict(X_validation)
print(accuracy_score(Y_validation, predictions))

0.8695652173913043
```

### G. Reporte de clasificación

El reporte obtenido nos muestra una puntuación-f1 (medida de exactitud de una prueba) de 0.88, lo cual indica que la precisión de la prueba realizada es alta.

Figura 33: Tabla de reporte de clasificación

		precision	recall	f1-score	support
<b>Reporte Resultados</b>	0	0.67	0.67	0.67	3
	1	1.00	1.00	1.00	7
	2	1.00	0.88	0.93	8
	3	0.33	0.50	0.40	2
	4	1.00	1.00	1.00	3
avg / total		0.90	0.87	0.88	23

### H. Clasificación de nuevos valores

Figura 34: Predicción de nuevos valores

```
X_new = pd.DataFrame({'a': [0], 'b': [14], 'c': [3], 'd': [2], 'e': [8], 'f': [1], 'g': [2]})
model.predict(X_new)

array([2], dtype=int64) ← Predicción
```

### I. Resultado de datos extraídos

A continuación, se muestran los resultados de la extracción de datos de las imágenes etiquetadas de emociones. Se muestra únicamente el resultado de una imagen por cada emoción, para facilidad de lectura.

Figura 35: Data final extraída y normalizada (Imagen etiquetada como alegría)

Parte	x	y	emotion	final
nose	0.4885969598619740	0.6307681025943420	Happy	noseHH,Happy
leftEye	1.0000007083316800	0.1415474447028840	Happy	leftEyeHH,Happy
rightEye	0.0000000000000000	0.0000000000000000	Happy	rightEyeHH,Happy
leftEar	1.7290963018265400	0.4995355660783800	Happy	leftEarHH,Happy
rightEar	-0.3759133173221550	0.3973878058586140	Happy	rightEarGH,Happy
leftShoulder	2.6823617355514900	2.2029158912765800	Happy	leftShoulderII,Happy
rightShoulder	-1.3374458830736500	2.0175507987737000	Happy	rightShoulderGI,Happy
leftElbow	5.0890381019974900	-0.3293562175826700	Happy	leftElbowJG,Happy
rightElbow	-3.3947962900555800	-0.7995988825260290	Happy	rightElbowFG,Happy
leftWrist	4.1521654161161500	-1.0350037069321400	Happy	leftWristJG,Happy
rightWrist	-0.9846235207222120	-1.9490451235526600	Happy	rightWristGG,Happy
leftHip	1.5680708128401900	5.5668256024859400	Happy	leftHipHJ,Happy
rightHip	-1.4522434239927400	5.5276624032005900	Happy	rightHipGJ,Happy
leftKnee	-4.6900768061817100	4.6240772499418600	Happy	leftKneeEJ,Happy
rightKnee	-4.5282193424926000	4.5863727110075900	Happy	rightKneeEJ,Happy
leftAnkle	-5.2055956232034200	4.6762942534959000	Happy	leftAnkleEJ,Happy
rightAnkle	-4.7184717841483500	4.6063348307418600	Happy	rightAnkleEJ,Happy

Figura 36: Datos finales extraídos y normalizados (Imagen etiquetada como miedo)

Parte	x	y	emotion	final
nose	0.327401643335	0.415938545212	fear	noseHH,fear
leftEye	1.000000806579	-0.139180778596	fear	leftEyeHG,fear
rightEye	0.000000000000	0.000000000000	fear	rightEyeHH,fear
leftEar	1.539142919750	0.250284448099	fear	leftEarHH,fear
rightEar	-0.921401550089	0.029757087415	fear	rightEarGH,fear
leftShoulder	2.591810703215	2.032699892186	fear	leftShoulderII,fear
rightShoulder	-1.172804379438	2.220107515302	fear	rightShoulderGI,fear
leftElbow	4.071763896183	4.670366289099	fear	leftElbowJJ,fear
rightElbow	-3.167157805824	5.109044568387	fear	rightElbowFJ,fear
leftWrist	2.394638803689	2.432790616229	fear	leftWristII,fear
rightWrist	-1.310868317020	2.358987775558	fear	rightWristGI,fear
leftHip	2.180410117597	8.244508243573	fear	leftHipL,fear
rightHip	-0.776252128135	8.698443311802	fear	rightHipGL,fear
leftKnee	1.526703441701	8.916454937654	fear	leftKneeHL,fear
rightKnee	0.672939048455	8.865135178392	fear	rightKneeHL,fear
leftAnkle	1.707497523465	5.044091819662	fear	leftAnkleHJ,fear
rightAnkle	-0.327334307585	3.644292785464	fear	rightAnkleGI,fear



Figura 37: Data final extraída y normalizada (Imagen etiquetada como tristeza)

Parte	x	y	emotion	final
nose	0.425826305036	0.412119321206	sad	noseHH,sad
leftEye	0.999998917695	-0.136204448450	sad	leftEyeHG,sad
rightEye	0.000000000000	0.000000000000	sad	rightEyeHH,sad
leftEar	2.655100872914	0.052534468101	sad	leftEarH,sad
rightEar	-0.417452431422	0.403099371523	sad	rightEarGH,sad
leftShoulder	4.793123560211	2.556801179675	sad	leftShoulderJI,sad
rightShoulder	-0.433230034755	3.073232952256	sad	rightShoulderGI,sad
leftElbow	6.006041709019	7.545170069496	sad	leftElbowKK,sad
rightElbow	-0.399716304512	8.474573595452	sad	rightElbowGL,sad
leftWrist	3.643115838933	5.691293511637	sad	leftWristJ,sad
rightWrist	1.274304064505	5.241581122210	sad	rightWristHJ,sad
leftHip	5.106711024705	8.749816536976	sad	leftHipJL,sad
rightHip	1.620380407986	8.686812714652	sad	rightHipHL,sad
leftKnee	5.977897810332	9.544045849069	sad	leftKneeJL,sad
rightKnee	-0.731318903465	9.574462355921	sad	rightKneeGL,sad
leftAnkle	4.706247052906	16.086067864986	sad	leftAnkleJ,sad
rightAnkle	-0.905632364610	16.780104164880	sad	rightAnkleG,sad

Figura 38: Data final extraída y normalizada (Imagen etiquetada como enojo)

Parte	x	y	emotion	final
nose	0.426837482783	0.873166442022	angry	noseHH,angry
leftEye	0.999997559708	0.171874491409	angry	leftEyeHH,angry
rightEye	0.000000000000	0.000000000000	angry	rightEyeGG,angry
leftEar	2.022500449296	0.362354712556	angry	leftEarH,angry
rightEar	-1.513175561392	0.321806042426	angry	rightEarGH,angry
leftShoulder	3.107522246799	4.571382129132	angry	leftShoulderIJ,angry
rightShoulder	-2.844657149269	4.227575467807	angry	rightShoulderFJ,angry
leftElbow	4.357890832164	8.843379548566	angry	leftElbowJL,angry
rightElbow	-4.963866993663	9.402432225499	angry	rightElbowEL,angry
leftWrist	3.585654349528	12.324630400512	angry	leftWristIN,angry
rightWrist	-4.817592948072	11.544063750510	angry	rightWristEM,angry
leftHip	1.596636672639	13.262379807920	angry	leftHipHN,angry
rightHip	-2.316834580087	13.082283075465	angry	rightHipFN,angry
leftKnee	3.071224342283	13.644463963278	angry	leftKneeIN,angry
rightKnee	-3.738152182455	13.447682534706	angry	rightKneeFN,angry
leftAnkle	3.396805960841	7.273107258515	angry	leftAnkleIK,angry
rightAnkle	-3.711927076360	11.289246885136	angry	rightAnkleFM,angry

Figura 39: Data final extraída y normalizada (Imagen etiquetada como sorpresa)

Parte	x	y	emotion	final
nose	0.496994104515	0.742281834390	surprise	noseHH,surprise
leftEye	0.999999119930	0.083404515697	surprise	leftEyeHH,surprise
rightEye	0.000000000000	0.000000000000	surprise	rightEyeGG,surprise
leftEar	1.273768927660	0.047130332638	surprise	leftEarHH,surprise
rightEar	-1.072412050349	-0.247536730198	surprise	rightEarGG,surprise
leftShoulder	2.197290257729	1.754994576529	surprise	leftShoulderIH,surprise
rightShoulder	-1.908759675097	2.278503480381	surprise	rightShoulderGI,surprise
leftElbow	3.781641376438	-0.984576082436	surprise	leftElbowIG,surprise
rightElbow	-2.221665187372	6.306397047521	surprise	rightElbowFK,surprise
leftWrist	1.331136797070	6.036026304417	surprise	leftWristHK,surprise
rightWrist	-2.250867899763	6.445591525974	surprise	rightWristFK,surprise
leftHip	2.770940497813	8.835186476517	surprise	leftHipIL,surprise
rightHip	-0.690324916858	9.343200434581	surprise	rightHipGL,surprise
leftKnee	5.935184600639	7.949258999583	surprise	leftKneeJK,surprise
rightKnee	4.735908795218	8.604766078339	surprise	rightKneeJL,surprise
leftAnkle	1.871811881147	9.370364262544	surprise	leftAnkleHL,surprise
rightAnkle	4.887508744316	8.633942872607	surprise	rightAnkleJL,surprise

Figura 40: Data final extraída y normalizada (Imagen etiquetada como disgusto)

Parte	x	y	emotion	final
nose	0.431970428627	0.559379398812	disgust	noseHH,disgust
leftEye	1.000000048073	0.007951251176	disgust	leftEyeHH,disgust
rightEye	0.000000000000	0.000000000000	disgust	rightEyeGHGH,disgust
leftEar	1.980270421213	0.420509397973	disgust	leftEarHH,disgust
rightEar	-0.377304985561	0.261244473740	disgust	rightEarGH,disgust
leftShoulder	2.938334490045	3.213951373407	disgust	leftShoulderII,disgust
rightShoulder	-0.073838254787	2.949224303667	disgust	rightShoulderGI,disgust
leftElbow	3.924810104051	7.536954533958	disgust	leftElbowIK,disgust
rightElbow	-1.636438562456	6.970263320252	disgust	rightElbowGK,disgust
leftWrist	3.027911077271	4.434721782035	disgust	leftWristIJ,disgust
rightWrist	3.161688823227	4.333787706367	disgust	rightWristIJ,disgust
leftHip	-0.424464861329	11.375770965895	disgust	leftHipGM,disgust
rightHip	3.510608990242	11.482523642695	disgust	rightHipIM,disgust
leftKnee	6.578622125633	7.519437462902	disgust	leftKneeKK,disgust
rightKnee	1.519116080977	11.227288121209	disgust	rightKneeHM,disgust
leftAnkle	1.466793755657	11.486901045920	disgust	leftAnkleHM,disgust
rightAnkle	1.423971031051	11.429612756512	disgust	rightAnkleHM,disgust

## VII. ANÁLISIS DE RESULTADOS

Cómo primera salida del algoritmo, se obtiene un arreglo de predicciones. Después de ejecutar el ajuste del modelo, se obtuvieron las predicciones del modelo, que son clasificaciones de cada elemento de los datos de las emociones. Las predicciones obtenidas son:

Figura 41: Predicciones obtenidas

```
[0 0 3 0 2 0 0 0 0 0 0 2 0 0 0 2 2 2 3 2 2 2 2 2 2 0 2 2 2 2 2 3 3 3 3 3 2  
0 2 0 2 3 3 0 3 1 1 1 1 1 1 1 1 4 1 1 4 4 1 4 4 4 1 4 4 4 4 1 4 4 4 4 1  
4]
```

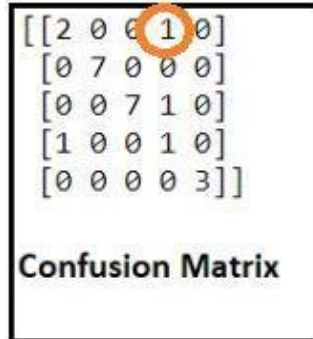
Estas predicciones obtuvieron un puntaje de precisión del 77.3%. Con este valor se confirma que el modelo fue bueno. Un valor cercano al 100% nos hubiera indicado un posible sobreajuste de los datos, es decir, que el modelo solo funciona para imágenes muy similares a las imágenes utilizadas en el entrenamiento. Un valor cercano al 50% nos hubiera indicado un subajuste, es decir, que el modelo no es lo suficientemente capaz para detectar nuevas imágenes con un alto grado de precisión. En el caso del resultado obtenido, un 77% indica que no existe subajuste ni sobreajuste y que el grado de precisión al predecir nuevas imágenes es bueno.

También se debe recalcar que el resultado de este valor depende de la cantidad de datos que se tenga para base de datos como entrada del algoritmo de regresión logística. Mientras más datos se tengan, mejor será el resultado (siempre y cuando las imágenes etiquetadas estén en su mayoría correctamente etiquetadas)

La matriz de confusión nos sirve para ver cuántos resultados equivocados se obtuvieron de cada clase. Es decir, cuántos resultados se predijeron que eran de una clase, cuando realmente eran de otra. Esto se ve en los valores que no forman parte de la diagonal. Las filas indican el valor de la categoría de donde el dato debería haber sido clasificado y las columnas indican la categoría a la que realmente fue clasificada (incorrectamente si son valores fuera de la diagonal de la matriz). En nuestro caso, los valores obtenidos nos indican lo siguiente:

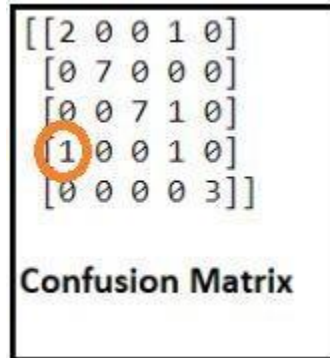
Existió una imagen de una persona feliz, incorrectamente clasificada por el algoritmo cómo enojada.

Figura 42: Persona feliz clasificada como enojada



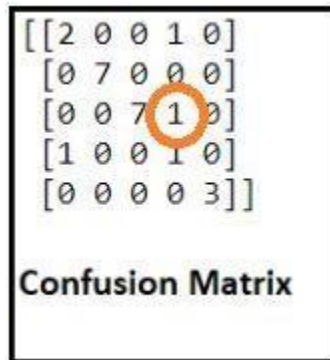
Existió una imagen de una persona enojada, incorrectamente clasificada como feliz.

Figura 43: Persona enojada clasificada como feliz



Existió una imagen de una persona que realmente estaba triste, incorrectamente clasificada por el algoritmo como enojada.

Figura 44 : Persona triste clasificada como enojada



Por otra parte, la precisión del modelo, calculada mediante el puntaje de precisión, se obtuvo de 86.95%. Este valor es bastante bueno, ya que nos indica que al comparar los conjuntos de validación cruzada y las predicciones hechas por el algoritmo, de 10 predicciones, 8.695 estuvieron correctas.

El resultado obtenido de f1-score, que nos indica precisión en una prueba, se obtuvo de 88%. Esto nos señala que la precisión de las predicciones es buena, y que existen muy pocos falsos positivos y muchos verdaderos positivos. Esto lo podemos observar más a detalle en la figura 33. En esta figura, podemos ver que el valor recall nos da 0.87. Este valor es el que indica la proporción de verdaderos positivos en todos los valores que efectivamente son positivos. Por otra parte, el valor de precisión se obtuvo de 0.90. El complemento de este valor nos indica cual es la proporción de falsos positivos en el resultado.

Para la parte de predicción de la emoción a partir de nuevas imágenes, se creó un arreglo de datos, que tiene cada punto clave de una imagen. Estos son los valores de cada parte del cuerpo, y en nuestro caso fueron numerados para la ejecución de la prueba de predicciones, con letras, comenzando con la letra “a”. En la figura número 34 se puede observar, que al ingresar los valores en todas las literales, en diferentes posiciones, se obtiene un valor numérico que corresponde a la emoción que representa dicho número.

En nuestro caso, podemos ver que el modelo generado a partir del conjunto de datos de prueba, predice la emoción de tristeza (representada por el número 2). Esto significa que, al notar la precisión obtenida del 88%, podemos decir que el modelo funciona bastante bien para la clasificación de las imágenes.

## VIII. CONCLUSIONES

- Se obtuvo un desempeño del modelo para el entrenamiento del 86% y un 77% de resultado de imágenes clasificadas correctamente.
- Se logró validar las imágenes etiquetadas del compilado, obteniendo una certeza de 92% sobre las imágenes etiquetadas personalmente vs las imágenes etiquetadas por la licenciada en psicología (15 imágenes)
- Se encontró que un compilado de 400 imágenes etiquetadas con las emociones básicas, es suficiente en tamaño para obtener buenos resultados, con una precisión alta de predicción de las imágenes.
- Se hizo efectiva la aplicación de los algoritmos de “aprendizaje de máquina”, detectando específicamente que el algoritmo de aprendizaje supervisado, denominado algoritmo de regresión logística, es apropiado para este tipo de problemas debido a los resultados obtenidos.
- Es necesaria una muestra grande de imágenes para mejorar la exactitud de los resultados, que de esta manera se evita que valores atípicos en la data alteren significativamente los resultados, o suceda sobre ajuste o sub ajuste con la data.
- Además de esto, se lograron normalizar los datos de todo el compilado de imágenes, llevando cada imagen a un mismo tamaño y posición, extrayendo la posición de cada punto clave del cuerpo y alimentando dichos datos al algoritmo de aprendizaje de máquina.

## IX. RECOMENDACIONES

- Al analizar una imagen etiquetada con una emoción, para obtener la estimación de la pose, se debe considerar que al usar el algoritmo para pose simple, este solo analiza una única persona, por lo que si en la imagen se muestra otra persona en el fondo, el algoritmo puede captar puntos clave incorrectos de la persona principal que se está analizando,
- En la parte de la discretización de las imágenes, se recomienda la utilización de una malla más grande, con más filas y columnas, para que las partes del cuerpo sean detectadas con mayor precisión.
- Para mejorar la exactitud del modelo generado, se recomienda la creación de una base de datos de imágenes de personas con más de 5000 fotos de personas, mostrando emociones.
- Respecto a las herramientas tecnológicas, aparte de la utilizada en esta investigación, se recomienda explorar la herramienta Open Pose, que detecta puntos clave de la mano, cara y pies. A su vez, se sugiere trabajar con personas del campo de la psicología, para incluir en el análisis de las emociones, las manos, gestos específicos de cara y pies, con el fin de realizar un proceso de etiquetado mucho más específico aún.

## X. REFERENCIAS BIBLIOGRÁFICAS

- Pease, A. A. (2017). *The Definitive Book of Body Language* (2004 ed.). Bantam.
- Ekman, P., & Friesen, W. V. (2003). *Unmasking the face: A guide to recognizing emotions from facial expressions*. Cambridge, MA: Malor Books.
- Avola D., Cinque L., Levialdi S., Placidi G. (2013) *Human Body Language Analysis: A Preliminary Study Based on Kinect Skeleton Tracking*. In: Petrosino A., Maddalena L., Pala P. (eds) *New Trends in Image Analysis and Processing – ICIAP 2013*.
- Vipin, K & Sharma, Vipin K. (2011). *Decoding Non-Verbal Communication*.
- Barkai, John & Ishikawa, Akira. (2018). *Nonverbal Communication from the Other Side: Speaking Body Language*.
- Bonoma, T., & Felder, L. (1977). *Nonverbal Communication in Marketing: Toward a Communicational Analysis*. *Journal of Marketing Research*, 14(2), 169-180. doi:10.2307/3150466
- Sikandar Amin, Mykhaylo Andriluka, and Marcus Rohrbach. (2013). *Multi-view Pictorial Structures for 3D Human Pose Estimation*.
- Mitchell, T. M. (2001). *Machine learning*. New York: McGraw-Hill.
- Berlatsky, N. (2011). *Artificial intelligence*. Detroit: Greenhaven Press.
- Russell, S. J., Norvig, P., & Canny, J. (2003). *Artificial intelligence: A modern approach*.
- Park, Hyeoun-Ae (2013). *An Introduction to Logistic Regression: From Basic Concepts to Interpretation with Particular Attention to Nursing Domain*.
- Darwin, C., Darwin, F., Barrett, P. and Freeman, R. (n.d.). *The expression of the emotions in man and animals*.
- Oved, Dan. *Real-time Human Pose Estimation in the Browser with TensorFlow.js* [En línea]. En: <https://medium.com/tensorflow/real-time-human-pose-estimation-in-the-browser-with-tensorflow-js-7dd0bc881cd5> Consultado el 02-11-2018
- Hurlin, Christopher. *Maximum Likelihood estimation*. [En línea]. En: [https://www.univ-orleans.fr/deg/masters/ESA/CH/Chapter2\\_MLE.pdf](https://www.univ-orleans.fr/deg/masters/ESA/CH/Chapter2_MLE.pdf) Consultado el 23-10-2018
- Assogba, Yannick. *Pose Detection in the browser: PoseNet Model*. [En línea] En: <https://github.com/tensorflow/tfjs-models/tree/master/posenet> Consultado el 23-10-



- Neverova, N & Kokkinos, I. *DensePose: Dense Human Pose Estimation In The Wild*. [En línea] . En: <https://github.com/facebookresearch/DensePose> Consultado el 29-10-2018
- Kim, Ildoo. *Tf Pose Estimation*. [En línea]. En: <https://github.com/ildoonet/tf-pose-estimation> Consultado el 15-11-2018

## XI. ANEXOS

A continuación se puede observar una fotografía de prueba de estimación de pose y posteriormente una parte del código de JavaScript utilizado para estimación de pose.

Figura 45: Estimación de pose, con imagen personal.

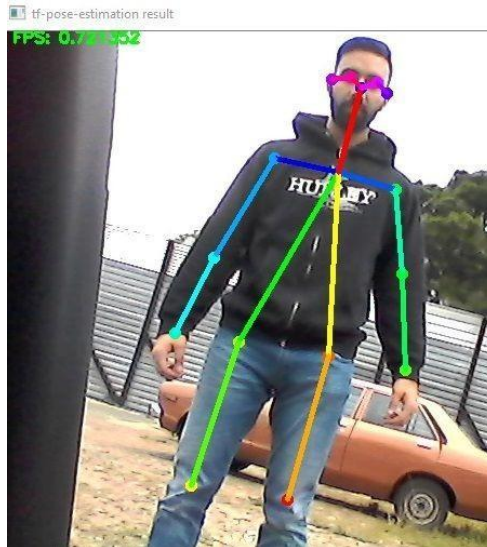


Figura 46: Código para estimación de pose

```
import * as posenet from '@tensorflow-models/posenet';
const imageScaleFactor = 0.3;
const outputStride = 8;
const flipHorizontal = false;

async function estimatePoseOnImage(imageElement) {
  // load the posenet model from a checkpoint
  const net = await posenet.load();

  const pose = await net.estimateSinglePose(imageElement, imageScaleFactor, flipHorizontal, outputStride);

  return pose;
}

let list = []

for (let i = 51; i < 70; i++) {
  list.push('surprise'+i);
}

let dataX = ""
for (let i = 0; i < 20; i++) {
  const result = estimatePoseOnImage(document.getElementById(list[i])).then(function(value) {
    for( let element of value.keypoints){
      dataX += `${element.position.x},${element.position.y},surprise `
    }
    dataX += `next`
    console.log(dataX)
    dataX = ""
  })
}
```