



UNIVERSIDAD POLITÉCNICA SALESIANA

SEDE QUITO

CARRERA DE COMPUTACIÓN

**VALIDACIÓN DE LA IDENTIFICACIÓN ÉTNICA MEDIANTE UN
SISTEMA DE VOTACIÓN POR MAYORÍA Y SU ANÁLISIS
ESTADÍSTICO EN LA POBLACIÓN DE ECUADOR**

Trabajo de titulación previo a la obtención del
Título de Ingenieros en Ciencias de la Computación

**AUTORES: ALEX DARIO TERREROS FREIRE
DIEGO GUILLERMO VICUÑA PILATAXI
TUTOR: DIEGO FERNANDO VALLEJO HUANGA**

Quito - Ecuador
2024

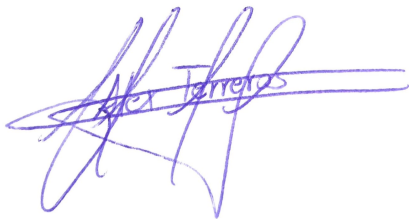
CERTIFICADO DE RESPONSABILIDAD Y AUTORÍA DEL TRABAJO DE TITULACIÓN

Nosotros, Alex Dario Terreros Freire con documento de identificación N.º 0706443157 y Diego Guillermo Vicuña Pilataxi con documento de identificación N.º 1721363974; manifestamos que:

Somos los autores y responsables del presente trabajo; y, autorizamos a que sin fines de lucro la Universidad Politécnica Salesiana pueda usar, difundir, reproducir o publicar de manera total o parcial el presente trabajo de titulación.

Quito, 04 de marzo del 2024

Atentamente,



Alex Dario Terreros Freire
0706443157



Diego Guillermo Vicuña Pilataxi
1721363974

CERTIFICADO DE CESIÓN DE DERECHOS DE AUTOR DEL TRABAJO DE TITULACIÓN A LA UNIVERSIDAD POLITÉCNICA SALESIANA

Nosotros, Alex Dario Terreros Freire con documento de identificación No. 0706443157 y Diego Guillermo Vicuña Pilataxi con documento de identificación No. 1721363974, expresamos nuestra voluntad y por medio del presente documento cedemos a la Universidad Politécnica Salesiana la titularidad sobre los derechos patrimoniales en virtud de que somos autores del Artículo Académico: Validación de la Identificación Étnica mediante un Sistema de Votación por Mayoría y su Análisis Estadístico en la Población de Ecuador, el cual ha sido desarrollado para optar por el título de: Ingenieros en Ciencias de la Computación, en la Universidad Politécnica Salesiana, quedando la Universidad facultada para ejercer plenamente los derechos cedidos anteriormente.

En concordancia con lo manifestado, suscribo este documento en el momento que hago la entrega del trabajo final en formato digital a la Biblioteca de la Universidad Politécnica Salesiana.

Quito, 04 de marzo del 2024

Atentamente,



Alex Dario Terreros Freire
0706443157



Diego Guillermo Vicuña Pilataxi
1721363974

CERTIFICADO DE DIRECCIÓN DEL TRABAJO DE TITULACIÓN

Yo, Diego Fernando Vallejo Huanga con documento de identificación N° 1720162708, docente de la Universidad Politécnica Salesiana, declaro que bajo mi tutoría fue desarrollado el trabajo de titulación: VALIDACIÓN DE LA IDENTIFICACIÓN ÉTNICA MEDIANTE UN SISTEMA DE VOTACIÓN POR MAYORÍA Y SU ANÁLISIS ESTADÍSTICO EN LA POBLACIÓN DE ECUADOR, realizado por Alex Dario Terreros Freire, con documento de identificación N.º 0706443157 y por Diego Guillermo Vicuña Pilataxi con documento de identificación N.º 1721363974, obteniendo como resultado final el trabajo de titulación bajo la opción Artículo Académico que cumple con todos los requisitos determinados por la Universidad Politécnica Salesiana.

Quito, 04 de marzo del 2024

Atentamente,



Ing. Diego Fernando Vallejo Huanga, MSc
1720162708

Validación de la Identificación Étnica mediante un Sistema de Votación por Mayoría y su Análisis Estadístico en la Población de Ecuador

1st Alex Dario Terreros Freire
aterrerosf@est.ups.edu.ec

2nd Diego Guillermo Vicuña Pilataxi
dvcunap@est.ups.edu.ec

3rd Diego Vallejo-Huanga
dvallejoh@ups.edu.ec

Resumen—Ecuador, como un Estado plurinacional y multiétnico, tiene una población diversa como producto de su mestizaje. Los grupos étnicos comparten rasgos y características físicas o genéticas, que los permiten identificar de manera unívoca. El proceso de identificación étnica es complejo porque depende de varios factores como culturales, geográficos, económicos y sociales. Una divergencia entre la auto-identificación y la identificación externa hacia un individuo podría generar conflictos económicos y sociales, en grupos étnicos ecuatorianos minoritarios, impidiéndoles ejercer su derecho a la auto-determinación y libre determinación. Esta investigación aborda la identificación étnica utilizando un conjunto de fotografías faciales de las cuatro etnias mayoritarias en Ecuador: afro-ecuatorianos, indígenas, euro-descendientes y mestizos. A partir de estas fotografías, se desarrolló un sistema web de votación por mayoría usando una estructura de datos llamada tabla hash para la recuperación de las imágenes y su posterior uso en el proceso de votación étnica. Con los datos obtenidos de las interacciones de usuarios, se realiza un análisis estadístico que mide la divergencia entre la auto-percepción y la percepción externa de la etnia en un individuo. El contraste de estos dos procesos de clasificación, se resumen en una matriz de confusión, donde el rendimiento medido con el *F1 score* es del 88.7% en la etnia afro-ecuatoriana, a diferencia de la etnia indígena con el 43.7%, mostrando así que las personas clasifican mejor a ciertas etnias, basados en una imagen facial. Por otro lado, en las ANOVAs se toma como referencia la prueba F, dando como resultado valores menores al nivel de significancia establecido en $\alpha = 0.05$. Con estos análisis estadísticos, se determina que no existe una diferencia significativa en términos de invariabilidad en la percepción de los grupos étnicos, lo cual proporciona una comprensión más profunda de la auto-percepción étnica.

Palabras Clave—Tabla hash, auto-identificación étnica, sistema de validación web, grupos étnicos ecuatorianos, ANOVA.

Abstract—As a plurinational and multiethnic State, Ecuador has a diverse population due to its miscegenation. Ethnic groups share physical or genetic traits and characteristics, which allow them to be uniquely identified. The ethnic identification process is complex because it depends on several factors, such as cultural, geographic, economic, and social. A divergence between self-identification and external identification towards an individual could generate economic and social conflicts in minority Ecuadorian ethnic groups, preventing them from exercising their right to self-determination. This research addresses ethnic identification using a set of facial photographs of the four majority ethnicities in Ecuador: Afro-Ecuadorian, indigenous, European-descended, and mestizos. From these photographs, a majority voting web system was developed using a data structure called a hash table

to recover the images and their use in the ethnic voting process. With the data obtained from user interactions, a statistical analysis is performed that measures the divergence between self-perception and external perception of an individual's ethnicity. The contrast of these two classification processes is summarized in a confusion matrix, where the performance measured with the *F1 score* is 88.7% in the Afro-Ecuadorian ethnic group, unlike the indigenous ethnic group with 43.7 %, thus showing that people classify certain ethnicities better, based on a facial image. On the other hand, in the ANOVAs, the F test is taken as a reference, resulting in values lower than the significance level established at $\alpha = 0.05$. These statistical analyses determine that there is no significant difference in invariability in the perception of ethnic groups, which provides a deeper understanding of ethnic self-perception.

Keywords—Hash table, ethnic self-identification, web validation system, Ecuadorian ethnic groups, ANOVA.

I. INTRODUCCIÓN

La forma en que una persona se reconoce étnicamente se sustenta en construcciones mentales de ideas, creencias, opiniones, actitudes y valores, creados por los miembros de una determinada etnia [1]. Las construcciones mentales están arraigadas en la historia, la cultura y las experiencias compartidas de un grupo étnico; estas influyen en la auto-percepción y en la identificación con esa etnia en particular. A lo largo del tiempo, estas pueden evolucionar y adaptarse, respondiendo a cambios en el contexto sociocultural y a la interacción con otras comunidades étnicas. Además, la identificación étnica no es estática, las personas pueden experimentar cambios en su auto-percepción étnica, influenciados por una variedad de factores como: la migración, la educación y la exposición a diferentes perspectivas culturales [2].

En la identificación étnica, a nivel mundial, se utiliza diferentes metodologías que analizan las costumbres y apropiaciones culturales como puntos clave para la auto-percepción étnica de un individuo [3]. En Indonesia, una investigación llevada a cabo durante el 2019 y 2020 señala que la identificación étnica de una persona también está correlacionada con su idioma de comunicación [4]. En Latinoamérica, un estudio realizado en Perú [5] con datos de 2012 a 2016 de la Encuesta Nacional de Hogares, investigó los rasgos individuales y locales relacionados con determinadas categorías de auto-identificación. Los resultados de la investigación muestran

que el número de personas que se identifican como indígenas varía según la pregunta utilizada. En otro estudio realizado en Bolivia, se emplea datos de encuestas entre 2004 y 2014 del Barómetro de las Américas de LAPOP, donde se expone el mismo resultado. La auto-percepción étnica varía de acuerdo a las preguntas planteadas, así, la población se percibía como indígena en un 62% para el 2001 y un 41% en el 2012. Por lo que se concluye que pequeños factores pueden intervenir en gran magnitud sobre la ideología y auto-percepción de una persona [6].

En el ámbito ecuatoriano, algunos pueblos y etnias no eran reconocidas por el Estado [7]. Por ejemplo, el pueblo montubio era considerado como mestizo hasta marzo de 2001, debido a que históricamente no tenían una taxonomía que les permita expresar su pertenencia hacia su grupo étnico. En el Censo de 2001, realizado por El Instituto Nacional de Estadísticas y Censos (INEC), los resultados de auto-identificación étnica muestran que el estado ecuatoriano se encuentra constituido por: 77.4% de mestizos, 10.5% de blancos, 6.8% de indígenas, 5.0% de afroecuatorianos y 0.3% de otras etnias. Luego, en el año 2011 el proceso censal emitió los siguientes resultados de auto-identificación étnica: mestizos 71.9%, indígenas 7.0%, afroecuatorianos 7.2%, blancos 6.1%, montubios 7.4% y otros 0.4% [8]. Estos antecedentes denotan que la identificación étnica de una persona depende de varios componentes y su percepción cambia a lo largo del tiempo.

La forma en la que se identifica la población étnicamente puede, en ciertos casos, producir conflictos entre los grupos minoritarios, quienes enfrentan sentimientos de auto-estigmatización como consecuencia de la presión social y estereotipos negativos asociados con su etnia [9]. Estos factores fomentan la fragmentación social entre grupos étnicos evitando la unidad y la inclusión, lo que puede dar lugar a la segregación y el aislamiento [10].

Considerando que el sentido de pertenencia étnica se modifica en el tiempo, debido a procesos de aprendizaje, donde la conformación de la identidad se afianza en distintos conflictos y tensiones que aparecen en su auto-percepción [11], el presente proyecto de investigación busca recolectar datos, mediante el uso de fotografías de distintos grupos étnicos, para conocer su percepción.

El objetivo principal de este trabajo de investigación es desarrollar un sistema de votación por mayoría con una interfaz web, permitiendo a los usuarios elegir el grupo étnico al que pertenecen las personas presentadas en las fotografías. Los grupos étnicos que se muestran están divididos en: eurodescendientes, indígenas, mestizos y afro-ecuatorianos. Los resultados del proceso de votación, mediante un análisis estadístico, medirán a-posteriori la discrepancia entre la auto-percepción y la percepción externa de la etnia en una persona. El análisis de la literatura científica, revela la falta de proyectos similares en Ecuador, por lo tanto, el desarrollo de este artículo es menester para procesos de investigación en identificación étnica y para conocer las dinámicas de cambio social. Además, se prevé que esta solución se mantenga operativa para futuros análisis y aplicaciones, incluyendo el desarrollo de algoritmos

de reconocimiento facial.

A. Trabajos Relacionados

Existen varias investigaciones y artículos relacionados con la identificación étnica, auto percepción e identificación externa de la identidad en varios contextos, ergo, es importante conocer algunos de los aportes más relevantes en este ámbito.

A nivel tecnológico, el estudio llevado a cabo por Lu et al. [12] emplea un conjunto de 2630 imágenes faciales humanas con el propósito de realizar una clasificación étnica entre individuos asiáticos y no asiáticos, mediante algoritmos de aprendizaje automático. Se utiliza diferentes resoluciones en las imágenes de entrada para la experimentación, aplicando un proceso de *Random Cross Validation* con 20 iteraciones. Los resultados revelan que el modelo *Latent Dirichlet Allocation* (LDA) a una resolución de 42x42 logra una precisión del 96%, superando al modelo KNN con una precisión del 94%, independientemente de la resolución empleada. Se concluye que, a pesar de la eficacia demostrada en este caso específico, la clasificación étnica basada en imágenes faciales no puede abordarse de manera exhaustiva mediante clasificadores lineales, debido a la complejidad inherente de la apariencia facial en contextos del mundo real. Como alternativa, se sugiere la consideración de clasificadores más avanzados, como las Máquinas de Soporte Vectorial (SVM), acompañadas de una representación más detallada de las características faciales, como puntos de referencia. Adicionalmente, los autores concluyen que los rasgos faciales son un factor importante para la percepción étnica de una persona.

Duan et al. [13] crearon una base de datos de rostros de grupos étnicos para extraer rasgos faciales, mediante el uso de tecnología de reconocimiento facial, dado que el rasgo facial es uno de los identificadores más importantes en una persona a nivel étnico. A través de un método de extracción de características, se obtienen dos tipos de funciones, algebraicas y geométricas, de la base de datos. Con el algoritmo LDA se extraen las funciones algebraicas de las imágenes y se construye una nueva plantilla facial. Por último, se utilizan los clasificadores KNN y C5.0 para generar modelos de aprendizaje automático. Los resultados indican que el promedio de precisión del modelo de reconocimiento en los grupos étnicos: Tibetano, Uigur y Zhuang pueden alcanzar el 79% por características algebraicas y el 90.95% por características geométricas.

Otro enfoque, dentro del proceso de auto-identificación étnica, se relaciona con los factores idiomáticos y culturales de las personas. Así, en el estudio de Batu et al. [4] realizado a 30 estudiantes sobre los idiomas que hablan e identidad personal y comunitaria, se afirma que 24 personas usan el indonesio como su primer idioma y solo tres afirmaron haber aprendido sus idiomas locales, Melayu, Balinés y Javanés, como lengua materna. Los resultados muestran que la población se siente étnicamente parte de Indonesia al hablar su idioma nativo e inglés, pero la otra sección de la población que habla idiomas locales siente confusión al momento de identificarse en un grupo étnico.

En el estudio de Rivas et al. [14] enfocado en las relaciones entre personas jóvenes que exploran su origen étnico, se analizó la relación entre la Identidad Étnico-Racial (IER) de los adolescentes y la diversidad demográfica de sus redes de amistad. La muestra consistió en adolescentes de diversos grupos étnico-raciales, 56% de afroamericanos, 31% de blancos y 13% de latinos en la región del medio oeste de los Estados Unidos, durante un año y seis meses, desde el 2014. Los resultados, derivados de modelos longitudinales cruzados, indicaron que los adolescentes que exploraron su IER en el verano de 2014 tendían a tener amistades más diversas después de seis meses y además mostraron una mayor exploración de su IER en el verano del 2015, un año después, independientemente de su grupo étnico. Con esta información, el autor concluye que la auto-percepción étnica está directamente vinculada con las relaciones sociales y sus interacciones.

En el caso particular de Ecuador, estudios muestran que en un contexto académico, jóvenes de etnias minoritarias pueden presentar dificultades para acceder a la educación superior. Estos grupos étnico minoritarios, generalmente, habitan en zonas rurales, ergo, tienen menos posibilidades de acceso a las universidades [15]. En este contexto, Grefa et al. [16] identifican las variables más latentes para que un joven acceda o no a la instrucción universitaria, mediante un modelo Logit. Con los datos recolectados, para la variable auto-identificación étnica, los resultados indican que, desde el año 2006 al 2014, se produce menor matriculación a la educación superior de los jóvenes de determinadas etnias. Etnias minoritarias como indígenas y afro-descendientes tienen menor acceso hacia la educación superior debido a disparidades culturales, económicas y sociales. Únicamente el 14.3% y 29.1%, respectivamente, de individuos pertenecientes a estas etnias tienen acceso a educación universitaria. Estas disparidades se reflejan en la calidad de los colegios, su infraestructura y la calidad de la enseñanza proporcionada por los docentes.

La divergencia entre la auto-percepción étnica frente a una identificación étnica externa, permite contextualizar ciertos problemas sociales que se presentan, mayoritariamente, en algunos grupos étnicos. Estudios han demostrado que existen grupos étnicos con mayores niveles de pobreza dentro de un mismo país. Estos grupos presentan escasez de acceso a agua potable y saneamiento adecuado y una limitada disponibilidad de alimentos nutritivos. Ortiz et al. [17] explican que el 25% de los niños menores de cinco años en Ecuador padecen desnutrición crónica. Esta problemática es más común en niños de origen indígena, afectando al 50% de esta población que, adicionalmente, tienen un 40% de algún grado de anemia. El estudio, de enfoque cuantitativo, se centró en 785 niños de la sierra ecuatoriana. Para evaluar el estado nutricional, se empleó el sistema informático SGINUTRICIÓN del Ministerio de Salud Pública. Los resultados indican una prevalencia del 18.1% en desnutrición según la relación peso-talla, siendo la baja estatura la condición más común con un 12.8%. Al considerar la talla en relación con la edad, el 29.5% de los niños mostró algún grado de desnutrición. De igual manera,

existe una diferencia significativa en la desnutrición según peso y talla entre niños indígenas, con un 35%, y mestizos con el 17%. De esta manera se concluye que la desnutrición crónica, medida por la relación peso-talla, supera en frecuencia a la desnutrición aguda, destacando diferencias significativas en la desnutrición entre mestizos e indígenas según la auto-identificación étnica.

II. METODOLOGÍA Y MATERIALES

El proceso metodológico para medir la divergencia entre la auto-identificación étnica y la percepción de un usuario externo, utiliza un instrumento de votación web desarrollado con herramientas de software libre. Este sistema de votación permite escoger, por mayoría, la etnia de un determinado individuo en función de las características faciales presentadas en una fotografía. El esquema de la Figura 1, muestra de forma resumida, la metodología que se utilizará para el proceso de validación de la identificación étnica propuesto en este trabajo de investigación.

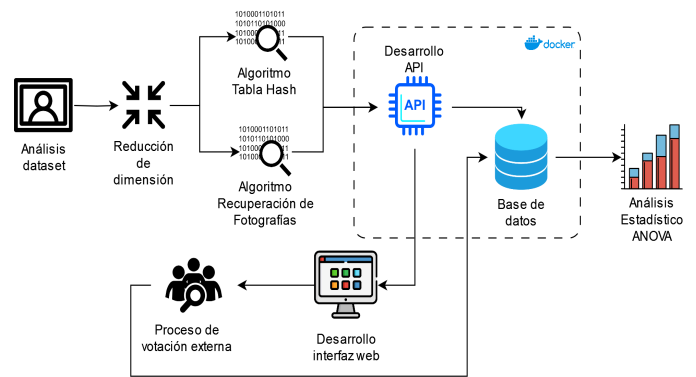


Fig. 1: Metodología de desarrollo del sistema de votación por mayoría

A. Descripción del Conjunto de Datos de Imágenes Faciales

Las fotografías que se muestran en el sistema web de votación fueron recolectadas en una investigación del año 2019, donde se recopilaban imágenes faciales de las cuatro etnias más prevalentes en Ecuador [18]. El conjunto de datos contiene un total de 430 imágenes, tomadas en cuatro ciudades del Ecuador, y cuyas taxonomías son: Mestizos, Afro-ecuatorianos, Euro-descendientes e Indígenas. Las imágenes tienen una dimensión de 1920x1280 en formato JPEG, adquiridas con una cámara Canon EOS Rebel T3i. Además, las fotografías están estandarizadas, en el escenario, mediante un fondo blanco y una distancia persona-cámara de 1.10m. El dataset tiene la cardinalidad descrita en la Tabla I.

Durante el proceso de recolección de los datos se realizó una encuesta a las personas fotografiadas preguntando sobre su auto-percepción tanto en etnia, como en los porcentajes de mestizaje, si los hubiere. Estos datos fueron almacenados en un fichero CSV que contiene los atributos de edad, género, etnia de auto-identificación y porcentajes de pertenencia hacia las etnias afro-ecuatoriana, euro-descendiente o indígena. El

Etnia	Sexo		Instancias
	Hombres	Mujeres	
Afro-ecuatorianos	39	11	50
Euro-descendientes	41	9	50
Indígenas	39	11	50
Mestizos	219	61	280
Total	338	92	430

Tabla I: Cardinalidad del número de instancias fotográficas distribuidas por etnia y sexo.

conjunto de datos considera que no solamente las personas mestizas pueden tener un porcentaje de afinidad hacia una de las tres principales etnias ecuatorianas, sino también aquellos individuos que se autoperceben, mayoritariamente, como afro-ecuatorianos, euro-descendientes o indígenas.

B. Descripción del Sistema de Votación por Mayoría

Los sistemas de votación mayoritaria tienen algunas variaciones en términos de aplicación y metodología. Con la llegada de las nuevas tecnologías de información y el auge de las aplicaciones de análisis de datos, la información de tipo no estructurada, como audios, textos e imágenes pueden ser explotadas de forma automatizada y con niveles de exactitud más altos. En este contexto, los investigadores están explorando varias formas de combinar las decisiones de múltiples clasificadores para ofrecer un rendimiento más preciso y robusto en una amplia gama de aplicaciones. Así, aparecen algoritmos combinados de decisión, denominados ensamblados [19], que son la base para los sistemas de votación [20].

Para el desarrollo de esta investigación se ha diseñado un sistema de votación por mayoría que permite elegir la etnia de una persona, mediante una fotografía presentada. El sistema implementa varios algoritmos tanto en el *front-end* como en el *back-end* con la finalidad de optimizar el rendimiento del sistema. La recuperación informática de las fotografías presenta un desafío debido a los tiempos de búsqueda y lectura. La solución planteada en esta investigación es el uso de tablas hash, que son una estructura de datos que agrupa valores en un índice y permite ordenar y recuperar datos mediante un mapeo clave-valor. La complejidad temporal para buscar, insertar y eliminar es en promedio $O(1)$, lo que significa que el tiempo de operación permanece constante independientemente del tamaño de la entrada. Esto resuelve el problema de la búsqueda lineal, que es costosa y tiene una complejidad temporal de $O(n)$ [21].

C. Desarrollo del Back-end del Sistema Web de Votación

La primera etapa en el desarrollo del *back-end* consiste en definir el proceso para el almacenamiento de las fotografías en una tabla hash.

La Figura 2 muestra este proceso, cuya salida es un archivo binario con extensión *pickle*. Cada fotografía, F_k , es leída y transformada al sistema de numeración posicional *Base64*. Este sistema de codificación permite la transmisión en formato JSON hacia el *front-end*.

Para almacenar cada F_k en la tabla hash se diseñó un algoritmo que realiza, en primera instancia, un proceso de cálculo

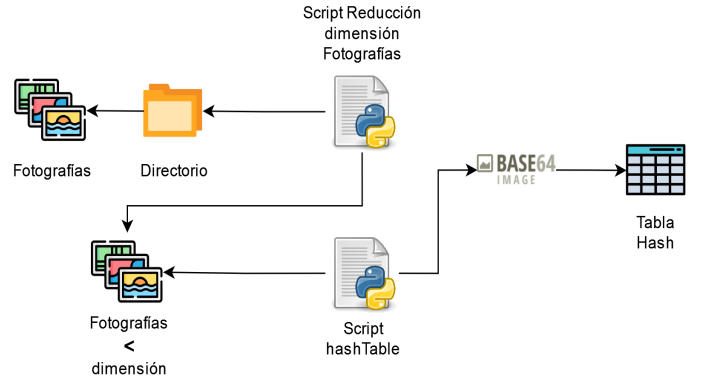


Fig. 2: Proceso de reducción de dimensión y almacenamiento de las fotografías en una tabla hash.

del índice para conocer la posición donde se almacenará la fotografía en el sistema de codificación *Base64*. La Ecuación 1 muestra el cálculo del índice i , donde k es la numeración de la fotografía y B es el número total de fotografías del *dataset*.

$$i = k \% B \quad (1)$$

Dado que en una tabla hash se pueden presentar colisiones, estas serán resueltas mediante direccionamiento abierto con estrategia de redispersión con segunda función hash. Entonces, cuando se detecta una colisión se usará la Ecuación 2 y Ecuación 3 para volver a calcular el índice considerando el número iteraciones C . Luego, el algoritmo podría presentar dos casos, cuando C es igual a 0 o mayor a 0. La Ecuación 2 muestra el primer caso, donde i_{old} es el índice anterior e i_{new} es el nuevo índice.

$$i_{new} = (i_{old} + k + C) \% B \quad (2)$$

Para el segundo caso $C > 0$ se utiliza la Ecuación 3, donde antes de calcular el nuevo índice i_{new} se ajusta el valor de k para evitar colisiones infinitas.

$$\begin{aligned} k_{new} &= (k \% (B - 1)) + 1 \\ i_{new} &= (i_{old} + k_{new} + C) \% B \end{aligned} \quad (3)$$

En el Pseudocódigo 1 se observa el procedimiento algorítmico completo para la generación de la tabla hash.

Para generar la tabla hash se realizaron dos pruebas usando un computador con las siguientes características: procesador Intel Core i7-8750H @2.20GHz, 16GB RAM y una tarjeta gráfica NVidia GTX1060. En la primera prueba se ejecutó el algoritmo con las fotografías en tamaño original, el proceso duró 8 minutos y generó un archivo binario con un tamaño de 480.5MB. El tamaño del fichero binario ocasionaba problemas de lectura y recuperación de las fotografías, por lo tanto, se redujo la dimensión de las fotografías en una escala de 1:2.5. Las nuevas fotografías tienen una dimensión de 768x512 y se utilizó un *script* en Python con la librería OpenCV, para este propósito. Luego, en la segunda prueba, se utilizó

Algoritmo 1 Tabla Hash

```
Entrada:  $B, k, F_k$ 
Salida:  $hash\_table$ 
1: Paso 1: Inicialización
2:  $C \leftarrow 0$ 
3: function  $hash(k)$ 
4:   return  $k \% B$ 
5: end function
6: function  $hash2(i, k)$ 
7:   if  $C > 0$  then
8:      $k_{new} \leftarrow (k \% (B - 1)) + 1$ 
9:   else
10:     $k_{new} \leftarrow k$ 
11:   end if
12:    $i_{new} \leftarrow (i + k_{new} + C) \% B$ 
13:   if  $colision(i_{new})$  then
14:      $table[i_{new}] \leftarrow F_k$ 
15:   else
16:      $C \leftarrow C + 1$ 
17:      $hash2(i_{new}, k_{new})$ 
18:   end if
19: end function
20: function  $colision(i)$ 
21:   if  $table[i] == NULL$  then
22:     return true
23:   else
24:     return false
25:   end if
26: end function
27: function  $hash\_table(k, F_k)$ 
28:    $i \leftarrow hash(k)$ 
29:   if  $colision(i)$  then
30:      $table[i] \leftarrow F_k$ 
31:   else
32:      $hash2(i, k)$ 
33:   end if
34: end function
```

las fotografías con menor dimensión y el proceso tardó 30 segundos, el tamaño del archivo fue de 35MB, teniendo como resultado la disminución del tamaño en un 92.71%.

El almacenamiento de las fotografías F_k en una tabla hash supone crear un algoritmo para la recuperación de las mismas. En el Pseudocódigo 2 se visualiza el proceso del algoritmo de recuperación de la información, cuyo funcionamiento se basa en la aleatorización de los índices iniciales y su posterior recuperación de las F_k . El algoritmo valida que, cuando el sistema de votación se ejecute en la web, no se muestre al usuario una fotografía repetida. El algoritmo tiene como salida un objeto JSON con cinco imágenes seleccionadas F_5 que se mostrarán al usuario.

Una *Application Programming Interface* (API) permite la comunicación entre el *back-end* y el *front-end* del sistema de votación a través de peticiones HTTP. En esta investigación se desarrolló una API en Python integrando el *framework* Flask que, adicionalmente, permite crear aplicaciones web. El envío de las fotografías y los datos de respuesta se realiza con el formato de texto para el intercambio de datos *JavaScript Object Notation* (JSON). La Figura 3 muestra el proceso que sigue el sistema cuando se inicia un proceso de votación por parte de un usuario. El *front-end* hace una petición GET y la API utiliza el algoritmo de recuperación de las F_k para buscar en el directorio de fotografías. Para el almacenamiento de las respuestas se hace una petición POST, que será procesada

Algoritmo 2 Recuperación de Fotografías

```
Salida:  $F_5$ 
1: Paso 1: Inicialización
2:  $F_5 \leftarrow []$ 
3:  $count \leftarrow 0$ 
4: Paso 2: Definición de funciones
5: function  $search\_hashtable(i)$ 
6:    $hash\_table \leftarrow open("hashtable.pickle")$ 
7:   return  $hash\_table[i]$ 
8: end function
9: function  $get\_images$ 
10:   $i \leftarrow randomInt(0, (k - 1))$ 
11:   $image \leftarrow \{\}$ 
12:   $response \leftarrow search\_hashtable(i)$ 
13:  if  $response$  then
14:     $image['data'] \leftarrow \{response[0], response[1], i\}$ 
15:  else
16:    return false
17:  end if
18: end function
19: while  $count < 5$  do
20:   $image = get\_images()$ 
21:  if  $image$  then:
22:     $F_5.add(image)$ 
23:     $count \leftarrow count + 1$ 
24:  end if
25: end while
```

por la API, y los datos se enviar a la base de datos para su almacenamiento.

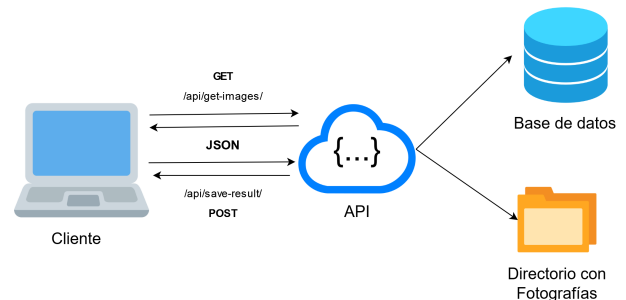


Fig. 3: Esquema de funcionamiento de la API

D. Desarrollo del Front-end del Sistema Web de Votación

Para el desarrollo del *front-end* se utilizó el *framework* VueJS 3 que permite el desarrollo de aplicaciones *Single Page Application* (SPA). Una SPA constituye un paradigma arquitectónico en el desarrollo de aplicaciones web que se caracteriza por la carga inicial de una única página HTML, seguida de la manipulación dinámica de su contenido mediante técnicas de navegación asíncrona [22]. Además, se utiliza la librería *Axios* para el procesamiento de las peticiones HTTP hacia la API y el *framework* *TailwindCSS* para los estilos web.

El sistema de votación por mayoría se desplegó en dos contenedores *Docker* utilizando el servicio en la nube *Google Cloud Platform* (GCP). El primer contenedor almacena la base de datos *PostgreSQL* y el segundo a la API desarrollada en Python. Para el *front-end* se utilizó un servicio de *hosting* gratuito llamado *InfinityFree*. La aplicación web desarrollada para la votación étnica está disponible en el enlace: <http://>

//ethnicsense.great-site.net/. La Figura 4 muestra el diagrama de bloques genérico del desarrollo *full-stack* de la aplicación web.

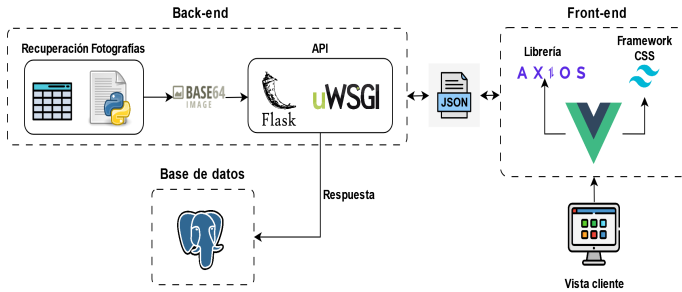


Fig. 4: Diagrama de bloques del desarrollo *full-stack* de la aplicación web

El sistema web cuenta con una interfaz de dos páginas, solo con elementos esenciales de funcionamiento, para facilitar el proceso de votación por parte de los usuarios. En la página principal se presentan dos botones uno para iniciar la votación y otro para leer más información acerca del proyecto de investigación. Al dar clic en iniciar *test* se muestran los pasos que el usuario debe seguir para completar la votación.

El proceso de votación tiene únicamente dos pasos, seleccionar una etnia y los porcentajes de mestizaje que considere el usuario en cada una de las fotografías. En el *front-end* se debe controlar el *scroll* de las barras de porcentaje, debido a que la suma de las tres barras no puede superar el 100%, como se visualiza en la Figura 5. Para validar esta condición se desarrolló un algoritmo que realiza el control con los datos que ingresa el usuario en tiempo real. El Pseudocódigo 3 muestra el proceso de validación que tiene como entrada los tres valores numéricos de los *scrolls*, $R1$ para la etnia afroecuatorialiana, $R2$ para la etnia euro-descendiente, $R3$ para la etnia indígena y un Id que es el identificador de cada *scroll*.



Fig. 5: Interfaz del sistema web de votación por mayoría.

E. Técnicas de Análisis Estadístico de los Datos

En esta investigación se usa una herramienta de análisis estadístico llamada *Analysis of Variance* (ANOVA), para medir la diferencia de las medias de varios grupos independientes en una observación. ANOVA utiliza la prueba F, desarrollada por Ronald Fisher, en la década de 1920, que evalúa diferencias potenciales en una variable dependiente de nivel continuo mediante una variable independiente de nivel categórico [23]. El método presupone independencia de muestras,

Algoritmo 3 Control de *scroll* en barras de porcentaje

Entrada: $Id, R1, R2, R3$

Salida: $[R1, R2, R3]$

```

1: Paso 1: Inicialización
2:  $R1 \leftarrow 33$ 
3:  $R2 \leftarrow 33$ 
4:  $R3 \leftarrow 34$ 
5: Paso 2: Definición de la función
6: function scroll_percentages( $R1, R2, R3$ )
7:    $a \leftarrow 100 - R1$ 
8:   if  $R2 == 33$  AND  $R3 == 34$  then
9:     if  $a \text{ MOD } 2 == 0$  then
10:       $R2 \leftarrow a/2$ 
11:       $R3 \leftarrow a/2$ 
12:     else
13:       $R2 \leftarrow (a - 1)/2$ 
14:       $R3 \leftarrow (a - 1)/2 + 1$ 
15:     end if
16:   else
17:      $b \leftarrow R1 + R2$ 
18:      $c \leftarrow 100 - b$ 
19:     if  $c > 0$  then
20:        $R3 \leftarrow c$ 
21:     else
22:        $R3 \leftarrow 0$ 
23:     end if
24:     if  $b > 100$  then
25:        $R2 \leftarrow 100 - R1$ 
26:     end if
27:   end if
28:   return  $[R1, R2, R3]$ 
29: end function
30: if  $Id == 1$  then
31:   scroll_percentages( $R1, R2, R3$ )
32: else if  $Id == 2$  then
33:   scroll_percentages( $R2, R1, R3$ )
34: else if  $Id == 3$  then
35:   scroll_percentages( $R3, R1, R2$ )
36: end if

```

distribuciones normales y homogeneidad de varianzas, y cuando se detecta una diferencia significativa, se recurre a pruebas *post hoc*, como la prueba de Tukey o la prueba de Bonferroni, para identificar diferencias específicas entre grupos [24]. En una ANOVA se plantean dos hipótesis que son:

H_0 : No hay diferencia entre los grupos o medias

H_a : Hay diferencias entre los grupos y medias

La prueba Tukey *Honestly Significant Difference* (HSD), constituye una herramienta estadística para evaluar la significancia entre dos conjuntos de datos y probar una hipótesis (H_0 o H_a). Su aplicación se justifica cuando la interacción entre tres o más variables es estadísticamente significativa. Utilizada en ANOVA, el método de Tukey genera intervalos de confianza para todas las comparaciones por pares entre las medias de los niveles de factores, controlando la tasa de error. Además, arroja el valor de probabilidad (*p-value*) asociado con cada comparación donde se indica la probabilidad de obtener una diferencia observada entre las medias de los grupos, si la hipótesis nula es verdadera. Un *p-value* pequeño sugiere que hay una diferencia significativa. La prueba de Tukey calcula y ajusta el nivel de confianza para cada comparación individual,

reduciendo así la tasa de error, y se ejecuta como prueba post hoc después de la recopilación de datos. El valor de la prueba se obtiene mediante la diferencia absoluta entre pares de medias dividida por el error estándar de la media, determinado por una prueba ANOVA unidireccional, siendo el error estándar la raíz cuadrada de la varianza dividida por el tamaño de la muestra [25].

Los grupos étnicos, G_j , se denotarán de la siguiente manera: afro-ecuatorianos G_1 , euro-descendientes G_2 , indígenas G_3 y mestizos G_4 . Entonces, el contraste de hipótesis para la ANOVA se define como:

$$H_0 : \mu_{G_1} = \mu_{G_2} = \mu_{G_3} = \mu_{G_4}$$

$$H_a : \mu_{G_1} \neq \mu_{G_2} \neq \mu_{G_3} \neq \mu_{G_4}$$

El nivel de confianza establecido para el rechazo de la hipótesis H_0 es del 95%, i.e., que el nivel de significancia $\alpha = 0.05$.

III. RESULTADOS Y DISCUSIÓN

El sistema web obtuvo al menos una respuesta de identificación étnica externa por cada fotografía, i.e., un total 430 respuestas. Estos datos de evaluación externa fueron comparados con la auto-identificación étnica obtenida a través de encuestas en el *dataset* [18]. Esta comparación dio como resultado los valores resumidos en la matriz de confusión de la Tabla III.

Auto-Identificación	Identificación Externa			
	G_1	G_2	G_3	G_4
G_1	47	0	2	7
G_2	0	28	0	33
G_3	2	1	31	58
G_4	1	21	17	182

Tabla II: Matriz de confusión entre datos de auto-identificación e identificación externa.

Considerando que la clase real es la auto-identificación étnica, desde la matriz de confusión, se pueden obtener algunas métricas para cuantificar cuan acertada es la identificación étnica externa. El *F1 score* es una métrica útil cuando se tiene datos desbalanceados debido a que combina los valores de *precision* y *recall*. Así, los valores obtenidos con esta métrica son de 0.887 para G_1 , 0.505 en G_2 , 0.437 para el G_3 y 0.727 en G_4 .

Los resultados determinan que los ecuatorianos identifican mejor a los afro-ecuatorianos y mestizos, en contra parte con las etnias euro-descendiente e indígena, donde el *F1 score* es aproximadamente del 50%.

Dado que la mayor divergencia de los resultados se encuentran en la etnia indígena, las Figuras 6 y 7 muestran los diagramas de caja y bigotes de las cuatro etnias frente el valor porcentual del nivel de indigenismo en los 430 individuos. En el primer gráfico, correspondiente a la auto-identificación, se visualiza que los valores de la etnia indígena

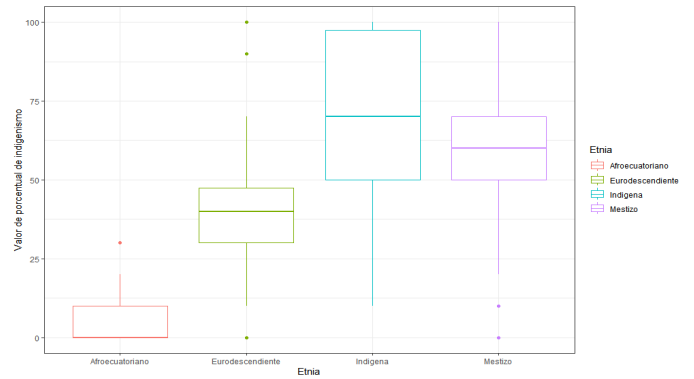


Fig. 6: Diagrama de caja y bigotes con los datos de auto-identificación del valor porcentual de indigenismo.

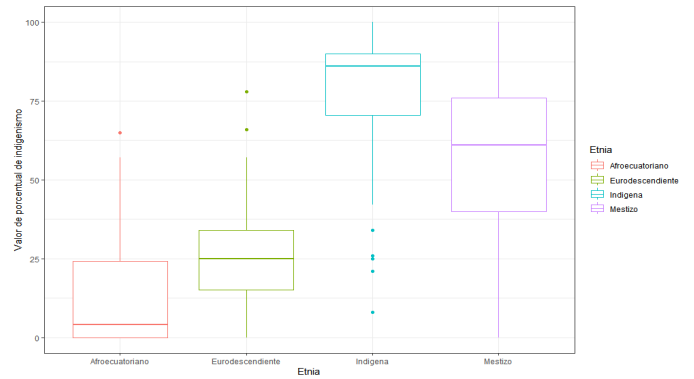


Fig. 7: Diagrama de caja y bigotes con los datos de identificación externa del valor porcentual de indigenismo.

tiene una alta variabilidad y una media similar a la etnia mestiza. Además, las personas que se auto-identifican como indígenas respondieron con porcentajes cercanos al 70% de correspondencia a esta etnia. En el segundo gráfico, la identificación externa muestra una variabilidad más alta en las etnias afro-ecuatoriana y mestiza, a diferencia de la etnia indígena que presenta una variabilidad menor. La media de percepción externa a individuos clasificados como indígenas supera el 80%, por lo tanto, las personas identifican con valores más altos a las personas con rasgos indígenas.

Tanto en el conjunto de datos del sistema web de Identificación externa (*IE*) como en el *dataset* de Auto-Identificación étnica (*AI*) se calculó la media y desviación estándar, para el posterior análisis de la varianza.

Los valores $(V_{i,j})$, que se pueden ver en la Tabla III, corresponden a los valores porcentuales de pertenencia de cada grupo étnico G_j . Este estudio considera que una persona no mestiza, puede tener un porcentaje de pertenencia, i , hacia otras etnias (afro-ecuatoriano, euro-descendiente e indígena) aunque con un porcentaje minoritario. Así, por ejemplo $V_{1,j}$ identifica el valor porcentual de la etnia afro-ecuatoriano, $V_{2,j}$ la etnia euro-descendiente y $V_{3,j}$ la etnia indígena. A partir de estos datos se genera cuatro ANOVAs para verificar las varianzas entre los valores numéricos porcentuales de cada

una de las etnias.

$V_{i,j}$	Dataset	μ	σ
$V_{1,1}$	AI	94.000	11.248
	IE	82.120	25.638
$V_{2,1}$	AI	2.000	5.345
	IE	5.340	10.513
$V_{3,1}$	AI	4.000	7.284
	IE	12.540	18.648
$V_{1,2}$	AI	4.000	10.302
	IE	8.100	13.579
$V_{2,2}$	AI	53.800	20.192
	IE	57.920	26.286
$V_{3,2}$	AI	42.200	20.433
	IE	33.980	23.822
$V_{1,3}$	AI	7.400	14.957
	IE	8.920	18.556
$V_{2,3}$	AI	21.000	17.642
	IE	16.580	19.149
$V_{3,3}$	AI	71.600	22.800
	IE	74.500	28.567
$V_{1,4}$	AI	6.143	13.039
	IE	10.200	15.989
$V_{2,4}$	AI	33.464	16.800
	IE	31.025	22.485
$V_{3,4}$	AI	60.393	16.467
	IE	58.775	24.750

Tabla III: Media y desviación estándar de cada valor diferenciado por cada grupo étnico y conjunto de datos.

Para la evaluación de las ANOVAs se toma como referencia la prueba F, donde se debe tener un nivel de significancia de $\alpha > 0.05$ para considerar que existe una variabilidad significativa entre los grupos. Los resultados de las ANOVAs fueron $p_value_{G_1} = 0.00137$, $p_value_{G_2} = 0.00272$, $p_value_{G_3} = 0.000174$ y $p_value_{G_4} = 0.000323$, por lo tanto, se llega a la conclusión que la diferencia entre las medias de los grupos no es lo suficientemente significativa en comparación con la variabilidad dentro de los grupos. En la Figura 8 se puede analizar los resultados de forma gráfica, mostrando que la variabilidad es significativa únicamente en cuatro de los grupos, $V_{1,1}$, $V_{3,3}$, $V_{3,4}$ y $V_{2,2}$. Por lo tanto, se rechaza la H_0 y se acepta H_a .

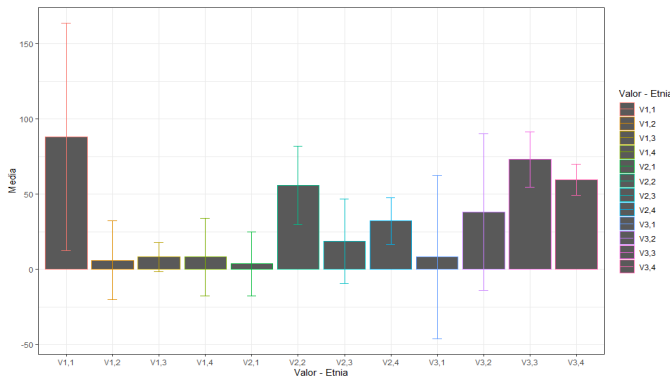


Fig. 8: ANOVA de las medias por Grupo étnico - $V_{i,j}$.

Considerando los resultados anteriores se puede ejecutar el *test* de Tukey, para comparar todas las combinaciones posibles de medias entre grupos y determinar cuáles de esas comparaciones son estadísticamente diferentes. La Tabla IV

Grupos	Diferencia	p_value
$V_{3,3} - V_{1,1}$	-15.010	0.069
$V_{1,3} - V_{1,2}$	2.110	0.999
$V_{1,4} - V_{1,2}$	2.121	0.999
$V_{2,1} - V_{1,2}$	-2.380	0.999
$V_{2,3} - V_{1,2}$	12.740	0.165
$V_{3,1} - V_{1,2}$	2.220	0.999
$V_{1,4} - V_{1,3}$	0.011	1.000
$V_{2,1} - V_{1,3}$	-4.490	0.986
$V_{3,1} - V_{1,3}$	0.110	1.000
$V_{2,1} - V_{1,4}$	-4.501	0.986
$V_{3,1} - V_{1,4}$	0.099	1.000
$V_{2,3} - V_{2,1}$	15.120	0.066
$V_{3,1} - V_{2,1}$	4.600	0.984
$V_{2,4} - V_{2,3}$	13.455	0.126
$V_{3,1} - V_{2,3}$	-10.520	0.353
$V_{3,2} - V_{2,4}$	5.845	0.925
$V_{3,4} - V_{3,3}$	-13.466	0.126

Tabla IV: Comparaciones no significativamente diferentes entre grupos con el *test* de Tukey.

muestra las comparaciones entre grupos donde se resumen solo las que tienen un valor de $p_value > 0.05$, indicando que la diferencia entre los grupos no es significativamente diferente.

IV. CONCLUSIONES

Esta investigación desarrolló un sistema web de votación por mayoría para obtener datos sobre la identificación étnica externa, permitiendo realizar un análisis estadístico y verificar la divergencia entre la auto-percepción y la percepción externa. Con este sistema se obtuvo al menos una respuesta por cada fotografía permitiendo generar una matriz de confusión, donde el rendimiento se mide con el *FI score* obteniendo valores de 0.887 para G_1 , 0.505 en G_2 , 0.437 para el G_3 y 0.727 en G_4 , demostrando que las personas clasifican mejor a los afro-ecuatorianos y mestizos, al contrario de las etnias euro-descendiente e indígena. Por otro lado, en las ANOVAs se toma como referencia la prueba F dando como resultado que ningún valor de p_value supera el nivel de significancia $\alpha = 0.05$. Ambos análisis estadísticos muestran que no existe una diferencia significativa entre la auto-identificación y la identificación étnica externa en la población de Ecuador. Además, los datos obtenidos del sistema web permiten cuantificar la precisión de la auto-identificación étnica, lo que proporciona información valiosa para futuros estudios y aplicaciones relacionadas con la identificación étnica en la población ecuatoriana.

El análisis estadístico arroja resultados en términos de invariabilidad significativa en la percepción de los grupos étnicos. Esto proporciona una comprensión más profunda de la dinámica de cambio social y la auto-percepción étnica en la población de Ecuador. Así mismo, la falta de proyectos similares en Ecuador resalta la importancia y pertinencia de este trabajo de investigación, no solo para su utilización en el campo social, sino también para futuros análisis y aplicaciones, incluyendo el desarrollo y mejora de algoritmos de reconocimiento facial.

Finalmente, como perspectivas futuras de esta investigación, se podría aumentar el número de fotografías usadas, así como

establecer una muestra más amplia para el sistema web de votación. De esta forma, sería posible realizar un análisis estadístico más exhaustivo y que correlacione más variables en el proceso de identificación étnica.

REFERENCES

- [1] I. Todorova-Pirgova, "Cultural images of the ethnic groups and ethnic interrelations in the balkans," *Ethnologies*, vol. 21, no. 2, pp. 147–175, 1999.
- [2] M. E. Varnum and I. Grossmann, "Cultural change: The how and the why," *Perspectives on Psychological Science*, vol. 12, no. 6, pp. 956–972, 2017.
- [3] L. Lenovský, "Identity as an instrument for interpreting the socio-cultural reality," *European Journal of Science and Theology*, vol. 11, no. 5, pp. 171–184, 2015.
- [4] L. Batu and P. Nancy, "Heritage language and ethnic identity: A study on students' ethnic identity and self-identification in jakarta," Ph.D. dissertation, SEKOLAH TINGGI ILMU PELAYARAN, 2020.
- [5] M. Moreno and M. Benavides, "Dynamics of ethnic and racial self-identification in contemporary peru," *Ethnic and Racial Studies*, vol. 42, no. 10, pp. 1686–1707, 2019.
- [6] D. E. Moreno Morales, "The mysterious case of the disappearing indians: changes in self-identification as indigenous in the latest intercensus period in bolivia," *Latin American and Caribbean Ethnic Studies*, vol. 14, no. 2, pp. 151–170, 2019.
- [7] M. Rodríguez-Cruz, "Construir la interculturalidad. políticas educativas, diversidad cultural y desigualdad en ecuador," *Íconos. Revista de Ciencias Sociales*, no. 60, pp. 217–236, 2018.
- [8] INEC, "Instituto nacional de estadística y censos." [Online]. Available: <https://www.ecuadorencifras.gob.ec/institucional/home/>
- [9] O. Bardol, S. Grot, H. Oh, E. Poulet, H. Zeroug-Vial, J. Brunelin, and E. Leaune, "Perceived ethnic discrimination as a risk factor for psychotic symptoms: a systematic review and meta-analysis," *Psychological Medicine*, vol. 50, no. 7, pp. 1077–1089, 2020.
- [10] B. Lancee, "Ethnic discrimination in hiring: comparing groups across contexts. results from a cross-national field experiment," pp. 1181–1200, 2021.
- [11] N. Zañartu Canihuante, A. Aravena Reyes, P. Grandón Fernández, F. Sáez Delgado, and C. Zañartu Canihuante, "Identidad étnica, discriminación percibida y procesos afectivos en jóvenes mapuches urbanos," *Cultura-hombre-sociedad*, vol. 27, no. 2, pp. 229–250, 2017.
- [12] X. Lu, A. K. Jain *et al.*, "Ethnicity identification from face images," in *Proceedings of SPIE*, vol. 5404. Citeseer, 2004, pp. 114–123.
- [13] X.-d. Duan, C.-r. Wang, X.-d. Liu, Z.-j. Li, J. Wu, and H.-l. Zhang, "Expression of concern: Ethnic features extraction and recognition of human faces," in *2010 2nd International Conference on Advanced Computer Control*, vol. 2. IEEE, 2010, pp. 125–130.
- [14] D. Rivas-Drake, A. J. Umaña-Taylor, D. R. Schaefer, and M. Medina, "Ethnic-racial identity and friendships in early adolescence," *Child Development*, vol. 88, no. 3, pp. 710–724, 2017.
- [15] D. Post, "Las reformas constitucionales en el ecuador y las oportunidades para el acceso a la educación superior desde 1950," *Education Policy Analysis Archives/Archivos Analíticos de Políticas Educativas*, vol. 19, pp. 1–24, 2011.
- [16] G. T. Grefa López and J. E. Pavón Navarrete, "Factores que inciden en el acceso a la educación de tercer nivel de los jóvenes en el ecuador," B.S. thesis, Quito: UCE, 2019.
- [17] M. C. M. Ortiz, M. R. R. Guamán, K. M. B. Moso, and J. I. Contreras, "Prevalencia de desnutrición infantil en menores de tres años en dos cantones de ecuador," *Revista Investigación En Salud Universidad De Boyacá*, vol. 8, no. 1, pp. 18–32, 2021.
- [18] J. Avilés, H. Toapanta, P. Morillo, and D. Vallejo-Huanga, "Dataset of Ethnic Facial Images of Ecuadorian People," 6 2019. [Online]. Available: https://figshare.com/articles/dataset/Dataset_of_Ethnic_facial_images_of_Ecuadorian_people/8266730
- [19] T. G. Dietterich *et al.*, "Ensemble learning," *The handbook of brain theory and neural networks*, vol. 2, no. 1, pp. 110–125, 2002.
- [20] A. F. R. Rahman, H. Alam, and M. C. Fairhurst, "Multiple classifier combination for character recognition: Revisiting the majority voting system and its variations," in *Document Analysis Systems V: 5th International Workshop, DAS 2002 Princeton, NJ, USA, August 19–21, 2002 Proceedings 5*. Springer, 2002, pp. 167–178.
- [21] E. Karimov, *Data Structures and Algorithms in Swift: Implement Stacks, Queues, Dictionaries, and Lists in Your Apps*. Apress, 2020.
- [22] E. A. Scott Jr, *SPA Design and Architecture: Understanding single-page web applications*. Simon and Schuster, 2015.
- [23] B. K. Das, D. N. Jha, S. K. Sahu, A. K. Yadav, R. K. Raman, and M. Kartikeyan, "Analysis of variance (anova) and design of experiments," in *Concept Building in Fisheries Data Analysis*. Springer, 2022, pp. 119–136.
- [24] T. K. Kim, "Understanding one-way anova using conceptual figures," *Korean journal of anesthesiology*, vol. 70, no. 1, pp. 22–26, 2017.
- [25] A. Nanda, B. B. Mohapatra, A. P. K. Mahapatra, A. P. K. Mahapatra, and A. P. K. Mahapatra, "Multiple comparison test by tukey's honestly significant difference (hsd): Do the confident level control type i error," *International Journal of Statistics and Applied Mathematics*, vol. 6, no. 1, pp. 59–65, 2021.