



**ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL  
FACULTAD DE INGENIERÍA EN ELECTRICIDAD Y COMPUTACIÓN**

**“SISTEMA DE BÚSQUEDA EN LÍNEA DE SITIOS TURÍSTICOS DENTRO  
DE LA COSTA ECUATORIANA UTILIZANDO UN MODELO DE  
PREDICCIÓN DE MINERÍA DE DATOS”**

**TESIS DE GRADO**

Previo a la obtención del Título de:

**INGENIERO EN COMPUTACIÓN ESPECIALIZACIÓN SISTEMAS  
TECNOLÓGICOS**

**INGENIERO EN COMPUTACIÓN ESPECIALIZACIÓN SISTEMAS  
TECNOLÓGICOS**

**INGENIERO EN COMPUTACIÓN ESPECIALIZACIÓN SISTEMAS  
TECNOLÓGICOS**

Presentado por:

Silvana Cristina Andrade Ramírez

Flor María Macías Bustamante

Iván Pedro Ochoa Brito

GUAYAQUIL - ECUADOR

2006- 2007

## **AGRADECIMIENTO**

En primer lugar agradecemos a Dios por guiarnos en nuestro camino, por ser quien nos ilumina siempre y porque nos acompaña en cada situación y está presente en nuestras decisiones.

A nuestros padres por ser los pilares de nuestra educación y desarrollo personal, gracias a ellos que siempre han confiado en cada uno de nosotros y están en cada uno de los momentos que más los necesitamos.

## DEDICATORIA

Cuando nací mi padre era el ser que aparecía para aplaudir mis últimos logros. Cuando me iba haciendo mayor, era una figura que me enseñaba la diferencia entre el bien y el mal. Durante mi adolescencia era la autoridad que me ponía límites a mis deseos. Ahora que soy adulta, sé que fue el mejor consejero y amigo que tuve.

Dedico la presente tesis a los seres más maravillosos que me dieron la vida y que me impulsaron en cada momento de mi vida. A cada una de las personas que se cruzaron en mi vida y fueron parte de mi crecer día con día y me enseñaron a través de las vicisitudes dar siempre lucha y ser un mejor ser humano.

Finalmente a mis compañeros, ahora amigos: Flor e Iván a quienes llegué a estimar mucho y con quienes compartí casi un año de mi vida.

Silvana

## DEDICATORIA

A mi familia por darme todo su apoyo incondicional, a mi padre por su enseñanza y amor, por su herencia: mi educación; a mi madre por su apoyo todos estos años, por su infinito amor y comprensión; a mis hermanos por permitirme contar con ellos siempre. A ti amor, por tus consejos y por animarme a seguir siempre adelante... Y a todas aquellas personas que de una u otra forma, colaboraron o participaron en la realización de esta tesis.

Flor

## DEDICATORIA

A mi familia por apoyarme en cada  
paso, a Dios por darme a mi familia.

A todos aquellos que en su paso por  
mi vida me ayudaron a continuar.

A lo que no creyeron que lo lograría,  
lo hice.

Iván

## TRIBUNAL DE GRADUACIÓN

---

Subdecano de la Facultad de  
Ingeniería en Electricidad y  
Computación  
Ing. Holger Cevallos

---

Director de Tópico  
Ing. Fabricio Echeverría Briones

---

Miembro del Tribunal  
Dra. Katherine Chiluiza

---

Miembro del Tribunal  
Ing. Marcelo Loor

## DECLARACIÓN EXPRESA

“La Responsabilidad del contenido de esta Tesis de Grado, nos corresponde exclusivamente; y el patrimonio intelectual de la misma a la ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL”.

(Reglamento de Graduación de la ESPOL).

---

Silvana Andrade R.

---

Flor María Macías B.

---

Iván Pedro Ochoa B.

## ÍNDICE GENERAL

	Pág.
<b>AGRADECIMIENTO .....</b>	<b>I</b>
<b>DEDICATORIA .....</b>	<b>II</b>
<b>DEDICATORIA .....</b>	<b>III</b>
<b>DEDICATORIA .....</b>	<b>IV</b>
<b>TRIBUNAL DE GRADUACIÓN.....</b>	<b>V</b>
<b>DECLARACIÓN EXPRESA.....</b>	<b>VI</b>
<b>ÍNDICE GENERAL.....</b>	<b>VII</b>
<b>ÍNDICE DE FIGURAS .....</b>	<b>X</b>
<b>ÍNDICE DE TABLAS.....</b>	<b>XI</b>
<b>INTRODUCCIÓN .....</b>	<b>12</b>
<b>CAPÍTULO 1.....</b>	<b>15</b>
<b>ANTECEDENTES Y JUSTIFICACIONES .....</b>	<b>15</b>
1.1. OBJETIVOS.....	16
1.2. ANTECEDENTES.....	17
1.3. TURISMO EN LA COSTA ECUATORIANA .....	19
1.3.1. TIPOS DE TURISMO EN LA COSTA ECUATORIANA .....	20
1.4. MEDIOS DE DIFUSIÓN EN LÍNEA DEL TURISMO DE LA COSTA ECUATORIANA.....	21
1.5. JUSTIFICACIÓN.....	23
1.6. ALCANCES .....	23
<b>CAPÍTULO 2.....</b>	<b>25</b>
<b>FUNDAMENTOS TEÓRICOS DE LA MINERÍA DE DATOS .....</b>	<b>25</b>
2.1. CONCEPTOS Y DEFINICIONES DE LA MINERÍA DE DATOS .....	26
2.2. TIPOS DE MODELOS DE MINERÍA DE DATOS.....	27
2.2.1. MODELO PREDICTIVO .....	27
2.2.2. MODELO DESCRIPTIVO.....	27
2.3. TAREAS Y TÉCNICAS DE LA MINERÍA DE DATOS .....	28
2.4. EL PROCESO DE EXTRACCIÓN DE CONOCIMIENTO .....	34
2.5. EL CÍRCULO VIRTUOSO DE LA MINERÍA DE DATOS .....	36
<b>CAPÍTULO 3.....</b>	<b>39</b>
<b>ANÁLISIS DEL SISTEMA.....</b>	<b>39</b>



3.1.	ANÁLISIS Y REQUERIMIENTOS DEL SISTEMA.....	40
3.1.1.	ALCANCE DE DATOS .....	43
3.1.2.	ACCESO DE LOS CLIENTES AL SISTEMA.....	45
3.2.	ANÁLISIS DE TÉCNICAS DE MINERÍA DE DATOS A IMPLEMENTARSE	46
3.3.	ANÁLISIS ORIENTADO A OBJETOS.....	49
3.3.1.	DIAGRAMA DE CASOS DE USO.....	49
3.3.2.	ESPECIFICACIONES DE CASOS DE USO .....	50
3.4.	ANÁLISIS DE LAS HERRAMIENTAS DE DESARROLLO .....	53
3.4.1.	CRITERIOS PARA LA ELECCIÓN DE LAS HERRAMIENTAS .....	53
3.4.2.	JUSTIFICACIÓN PARA LA ELECCIÓN DE LAS HERRAMIENTAS SELECCIONADAS.....	53
<b>CAPÍTULO 4.....</b>		<b>59</b>
<b>DISEÑO DEL SISTEMA.....</b>		<b>59</b>
4.1.	VISIÓN GENERAL DEL DISEÑO.....	60
4.2.	DISEÑO ORIENTADO A OBJETOS.....	61
4.2.1.	DIAGRAMAS DE INTERACCIÓN DE OBJETOS.....	62
4.3.	DISEÑO DE LA BASE DE DATOS .....	65
4.3.1.	MODELAMIENTO DE LA BASE DE DATOS .....	65
4.3.2.	DISEÑO MULTIDIMENSIONAL DE LA BASE DE DATOS .....	68
4.4.	DISEÑO DE LA APLICACIÓN Y LA INTERACCIÓN CON EL USUARIO....	72
<b>CAPÍTULO 5.....</b>		<b>76</b>
<b>IMPLEMENTACIÓN Y PRUEBAS DEL SISTEMA.....</b>		<b>76</b>
5.1.	IMPLEMENTACIÓN DE MÉTODOS DE MINERÍA DE DATOS.....	77
5.1.1.	IMPLEMENTACIÓN DE TÉCNICAS BAYESIANAS PARA EL MODELO PREDICTIVO .....	77
5.1.2.	IMPLEMENTACIÓN DE LA FUNCIÓN LINEAL DISCRIMINANTE PARA EL MODELO DESCRIPTIVO .....	82
5.2.	IMPLEMENTACION DEL SISTEMA.....	85
5.3.	RESULTADOS DE LA APLICACIÓN DE MÉTODOS DE MINERÍA DE DATOS .....	90
5.3.1.	RESULTADOS DE TÉCNICAS BAYESIANAS PARA EL MODELO PREDICTIVO .....	91
5.3.2.	RESULTADOS DE LA FUNCION LINEAL DISCRIMINANTE PARA EL MODELO DESCRIPTIVO.....	93
5.4.	PRUEBAS REALIZADAS .....	94
<b>CAPÍTULO 6.....</b>		<b>99</b>
<b>PROPUESTA COMERCIAL DEL SISTEMA .....</b>		<b>99</b>
6.1.	ANALISIS FODA.....	100
6.2.	PROPUESTA DEL VALOR COMERCIAL DEL SISTEMA.....	102
6.3.	EVALUACIÓN DE COSTOS .....	103
6.4.	REQUERIMIENTOS MÍNIMOS DEL SISTEMA .....	105
<b>CONCLUSIONES Y RECOMENDACIONES.....</b>		<b>107</b>
<b>REFERENCIAS BIBLIOGRÁFICAS .....</b>		<b>110</b>

<b>ANEXOS.....</b>	<b>112</b>
ANEXO 1.....	113
TURISMO EN EL ECUADOR .....	113
ANEXO 2.....	121
ENCUESTA REALIZADA PARA RECONOCER ATRIBUTOS DE ATRACTIVOS TURÍSTICOS.....	121
ANEXO 3.....	127
CLASIFICACIÓN DE ACTIVIDADES TURÍSTICAS.....	127
ANEXO 4.....	128
TIPOS DE ATRACTIVOS TURÍSTICOS .....	128
ANEXO 5.....	129
FLUJO DE VENTANAS .....	129
ANEXO 6.....	130
DIAGRAMA DE CLASES.....	130
ANEXO 7.....	131
ESCENARIOS DE LOS CASOS DE USO DEL SISTEMA.....	131
ANEXO 8.....	136
MANUAL DE USUARIO.....	136
ANEXO 9.....	160
DICCIONARIO DE DATOS.....	160
ANEXO 10.....	174
ANÁLISIS DE ALGORITMOS A IMPLEMENTARSE .....	174
ANEXO 10A .....	175
EJEMPLO DE APLICACIÓN DE ALGORITMO ID3.....	175
EJEMPLO DE APLICACIÓN DE ALGORITMO VECINOS MÁS PRÓXIMOS .....	190
ANEXO 10C .....	195
EJEMPLO DE APLICACIÓN DE ALGORITMO NAIVE BAYES .....	195
ANEXO 10D .....	200
IMPLEMENTACIÓN DE LA FUNCIÓN LINEAL DISCRIMINANTE .....	200
ANEXO 11.....	203
RESULTADOS DE LA IMPLEMENTACIÓN DEL ALGORITMO DE NAIVE BAYES.... .....	203
ANEXO 12.....	209
VECTORES DE PROBABILIDAD DE CADA ACTIVIDAD .....	209
ANEXO 13.....	218
MATRIZ DE VARIABLES USADAS PARA LA PREDICCIÓN .....	218

## ÍNDICE DE FIGURAS

	<b>Pág.</b>
Figura 2.1. Fases del proceso de extracción de conocimiento .....	34
Figura 2.2. Círculo virtuoso de la minería de datos .....	36
Figura 3.1. Diagrama de casos de uso .....	50
Figura 4.1. Arquitectura del Sistema.....	60
Figura 4.2. Sistema realiza búsqueda de sitios turísticos exitosamente.....	62
Figura 4.3. Sistema no puede realizar búsqueda de sitios turísticos exitosamente.....	63
Figura 4.4. Sistema actualiza calificación y ranking de sitios turísticos exitosamente. ....	63
Figura 4.5. Sistema no actualiza calificación y ranking de sitios turísticos exitosamente.....	64
Figura 4.6. Recomendar sitios turísticos exitosamente .....	64
Figura 4.7. No se puede recomendar sitios turísticos exitosamente .....	65
Figura 4.8. Diseño de la base de datos sitios turísticos.....	68
Figura 4.9. Diseño multidimensional de la base de datos.....	72
Figura 4.10. Esquema del diseño del sistema.....	74

## ÍNDICE DE TABLAS

	<b>Pág.</b>
Tabla 2.1. Tareas de la minería de datos.....	28
Tabla 2.2. Técnicas más usadas en la minería de datos.....	29
Tabla 2.3. Tareas y tipo de técnica que pueden resolver dicha tarea.....	33
Tabla 4.1 Entidades descriptivas del Modelo de la Base de Datos.....	66
Tabla 4.2 Relaciones entre entidades.....	66
Tabla 4.3 Relación entre entidades intermedias.....	67
Tabla 4.4 Tablas de Hecho.....	69
Tabla 4.5 Tablas de dimensión.....	69
Tabla 4.6 Relaciones entre entidades del modelo multidimensional.....	70
Tabla 4.7 Relación entre entidades intermedias.....	71
Tabla 5.1. Matriz de registros.....	79
Tabla 5.2. Vectores de probabilidades de cada variable.....	80
Tabla 5.3. Matriz de registros.....	81
Tabla 5.4. Vectores de probabilidades de cada variable.....	82
Tabla 5.5. Matriz de calificaciones de los atributos.....	84
Tabla 5.6. Coeficientes de probabilidad usados.....	84
Tabla 5.7. Usuarios de pruebas del sistema.....	95
Tabla 6.1. Software utilizado en el desarrollo del sistema.....	103
Tabla 6.2. Software y hardware recomendado para la implementación del sistema.....	104

## INTRODUCCIÓN

La presente tesis está basada en el desarrollo de un Sistema de búsqueda en línea de sitios turísticos utilizando un modelo de predicción, a través de este trabajo, se pretende mejorar la difusión de los sitios turísticos de la costa ecuatoriana, ofreciendo a los visitantes una forma de identificar cual es el lugar más adecuado a sus requerimientos.

En base de la información recopilada de actividades y características de algún atractivo, se genera el modelo de predicción utilizando una técnica bayesiana, ya que gracias a la implementación de esta técnica en el sistema, se logra estimar o predecir atractivos y sitios turísticos de la costa ecuatoriana, de tal manera que maximice la cobertura de exigencias en los usuarios y turistas en la búsqueda de sitios y atractivos.

En efecto, se espera realizar predicciones sobre los posibles atractivos que puede visitar el turista y que den con su perfil turístico, así como también mostrar una categorización de los mejores sitios en base a una calificación dada por el usuario que tenga conocimiento de los atractivos de la costa ecuatoriana.

El sistema incluye varios módulos diseñados para poder ser ejecutados de manera independiente por medio de un navegador web.

Para un mejor conocimiento del proyecto, se han documentado 5 capítulos, en los cuales se detallan cada una de las etapas del proyecto:

En el primer capítulo, se presentan los objetivos, justificaciones y alcances del sistema. Se revisan además los antecedentes, las diferentes formas que se ha difundido en el Internet el turismo de la costa ecuatoriana y los diferentes tipos y clasificaciones del turismo que se encuentran en la misma.

En el segundo capítulo, se explican las definiciones básicas en las cuales se fundamenta el desarrollo del proyecto. Se introducen: el concepto de minería de datos, las técnicas y modelos implementados, utilizando las fases del círculo virtuoso de la minería de datos.

En el tercer capítulo, se describe el análisis y requerimientos del sistema para llegar al modelo predictivo. Se muestra el diagrama de casos de uso, se explica cada uno de ellos junto con los actores que intervienen en el sistema.

El cuarto capítulo, detalla el diseño del sistema, la interacción con el usuario, modelo de base de datos, implementación de algoritmos de minería de datos, así como las herramientas tecnológicas para el desarrollo del mismo.

En el quinto capítulo, se realiza el análisis FODA del sistema, la propuesta del valor comercial del sistema y la evaluación de costos del sistema. Además se detalla todas las pruebas que conllevó la realización de cada una de las etapas del sistema para que pueda ser usado con total confianza por el usuario final, y que todos los requisitos se hayan cumplido. Finalmente se mencionan cuales son los requerimientos mínimos del sistema (hardware y software), necesarios para su funcionamiento.

## CAPÍTULO 1

# **ANTECEDENTES Y JUSTIFICACIONES**



## 1.1.OBJETIVOS

El objetivo general del sistema de búsqueda en línea de sitios turísticos de la costa ecuatoriana, es permitir a cualquier usuario, sin necesidad de registrarse previamente en el sistema, generar una predicción del lugar más adecuado que debe visitar, con eficacia, confiabilidad y sin costo alguno, utilizando una interfaz gráfica basada en Web, de fácil uso e interacción al usuario.

Los objetivos específicos que se desea conseguir con el desarrollo del sistema son:

- Desarrollar una herramienta que brinde servicio en línea exhibiendo los lugares turísticos de la costa ecuatoriana acorde a las necesidades y requerimientos del visitante.
- Identificar los sitios turísticos aptos a las preferencias del usuario, creando un modelo de predicción basado en su perfil turístico.
- Crear una base de conocimiento de sitios turísticos de la costa ecuatoriana aplicando métodos y tareas de la minería de datos.
- Generar predicciones que se acerquen a lo que el usuario desea conocer, utilizando como datos, las características y actividades del atractivo proporcionados por el visitante.

- Promover los lugares turísticos que existen en la costa ecuatoriana a nivel regional y mundial.

## 1.2. ANTECEDENTES

Internet y otras tecnologías informáticas están revolucionando la forma de trabajar de las empresas turísticas y el modo en que se promueven los destinos en el mundo.

En nuestro país, existen varios sistemas de difusión en línea del turismo, los cuales muestran información variada de las diferentes ciudades y sus atracciones.

Uno de estos ejemplos es el Ministerio de Turismo del Ecuador, el cual en su sitio Web [1], muestra información de las diferentes regiones del país e información acerca de las empresas que prestan servicio de transporte aéreo que tienen frecuencias hacia Ecuador.

También, encontramos sistemas como el *Sabre Travel Networks*, que permite hacer reservaciones en línea de hoteles, mediante diferentes sitios de Internet [2], otros sitios de renta de vehículos [3] y compra de boletos aéreos en línea [4].

En el 2005, Ecuador tuvo 860.784 visitantes extranjeros<sup>1</sup>, siendo los meses de: marzo, junio, julio y agosto; los de mayor flujo de visitantes. Los puertos de mayor ingreso de visitantes fueron los aeropuertos de Quito y Guayaquil con 345.400 y 187.270, respectivamente<sup>2</sup>.

Aunque en el Internet se encuentra una gran variedad de información de las actividades turísticas, ésta se encuentra muy dispersa en los diferentes sitios en línea, lo que dificulta a un visitante obtener toda la información necesaria para tomar la decisión al momento de viajar dentro del Ecuador.

Debido a los antecedentes mencionados, se hace necesario el desarrollo de un sistema que provea información concisa y detallada de las ciudades y atractivos turísticos junto con sus características y actividades, que se ajusten a las preferencias del usuario, permitiendo mediante el uso de la minería de datos realizar predicciones.

---

<sup>1</sup> Información proporcionada por el Departamento de Estadísticas del Ministerio de Turismo.

<sup>2</sup> Datos proporcionados por la Dirección Nacional de Migración.

### **1.3. TURISMO EN LA COSTA ECUATORIANA**

La costa ecuatoriana se encuentra bañada por el océano Pacífico, está conformada por cinco provincias, de norte a sur, que son: Esmeraldas, Manabí, Guayas, Los Ríos y El Oro.

La región se caracteriza por un clima cálido y seco al sur, y tropical húmedo al norte. Existen varios parques nacionales, reservas ecológicas y bosques; que ofrecen la posibilidad de realizar ecoturismo.

De norte a sur presenta hermosos balnearios con extraordinarias playas, arena limpia y brillante sol, que permiten al turista gozar de un clima agradable, de deportes acuáticos, de incomparable pesca y de toda clase de entretenimientos en confortables hoteles. Los balnearios del Ecuador brindan todas las comodidades de vivienda y turismo.

Mayor información acerca del turismo de la costa ecuatoriana se detalla en el anexo 1.

### 1.3.1. TIPOS DE TURISMO EN LA COSTA ECUATORIANA

De acuerdo a la intención que tenga el turista, o que se persiga al practicar el turismo, éste puede clasificarse de mil maneras. Pudiéramos llenar infinidad de páginas clasificando el turismo dentro de nuestras Costas.

Cualquiera que sean las causas por la que el hombre turísticamente se desplace, éstas van sugiriendo la clasificación del mismo. Tomemos algunas de ellas, pongámosle sus nombres y así iremos clasificando el turismo, según vayan aflorando: motivos, intenciones, deseos, oportunidades, necesidades y en fin tantas otras.

Se ha considerado utilizar, para efectos de desarrollo del sistema, una clasificación del turismo por las actividades que se puedan realizar; las cuales son:

- Turismo Deportivo: Es aquel en el cual las personas buscan lugares donde puedan practicar algún tipo de deporte.
- Turismo Cultural: Es donde la persona busca satisfacer su sed de conocimiento, buscando lugares que contengan teatros, museos, bibliotecas, sitios históricos y religiosos.

- Turismo Ecológico: Trata de acercar al turista con la naturaleza.
- Turismo Aventura: Viajes y salidas turísticas que implican un alto grado de contacto con la naturaleza mediante la acción y un cierto grado de riesgo implícito. Es un turismo que plantea desafíos y una sensación de riesgo (aunque las medidas de seguridad lo reduzcan a su mínima expresión) que el turista decide correr.
- Turismo Recreacional: Es el turismo de las masas en el cual los turistas tratan de encontrar lugares de relajamiento.

#### **1.4. MEDIOS DE DIFUSIÓN EN LÍNEA DEL TURISMO DE LA COSTA ECUATORIANA**

Internet, la autopista de la información como también se la denomina a esta red mundial de computadoras, se está constituyendo en unos de los medios más utilizados para acceder a fuentes de información en todo el mundo, en la más variada gama de actividades incluyendo el turismo [5].

En nuestro país existen diversas organizaciones privadas y publicadas que se dedican a difundir el turismo utilizando diversos medios de propagación de información turística. Instituciones como el Ministerio de Turismo, diarios, agencias de viajes, hoteles y todas las áreas

protegidas como: Reservas Forestales, Parques Nacionales Ecológicos, Reservas Faunísticas y Geobotánicas del País utilizan medios electrónicos (como el Internet), para promocionarse o informar de sus actividades al público en general.

Las empresas privadas del Ecuador que utilizan Internet para su beneficio institucional promueven al turismo en nuestro País directa o indirectamente.

La institución que difunde el turismo en nuestro País es el Ministerio de Turismo que tiene una página en Internet desde Junio de 1996, el cual se mantiene al día con las diferentes actividades turísticas que se realizan en el país y contiene información de los diferentes destinos turísticos de todo el país.

Existen también sitios Web de la mayoría de provincias de la costa y de las ciudades más importantes, pero en estos a veces la información referente al turismo es mínima, siendo su principal función comunicar políticas, resoluciones u ordenanzas.

También encontramos páginas Web creadas por personas que desean promocionar o comunicar sus localidades, que en general brindan

información sobre una ciudad o van dirigidos a su comunidad. En ciertos casos las páginas Web personales contienen información antigua, por el poco mantenimiento.

Ante todos estos medios de difusión en línea poco personalizados, el sistema implementado en esta tesis es una alternativa a los sistemas mencionados anteriormente que busca brindar información turística necesaria al usuario, visitante o turista para ayudarlo a tomar una decisión al momento de realizar turismo en las costas del Ecuador.

## **1.5. JUSTIFICACIÓN**

Debido al problema indicado anteriormente se hace necesario desarrollar un sistema que permita al visitante ingresar sus diferentes expectativas para un viaje y por medio de ellas mostrarle los diferentes lugares que sean aptos para su viaje.

## **1.6. ALCANCES**

- El sistema abarca únicamente los sitios turísticos de la costa ecuatoriana, enfocándose en sus características tales como: la limpieza, seguridad, concurrencia, lugar de parqueo y lugar para



comer; y sus actividades, que comprenden todas aquellas que el usuario/turista/visitante desee realizar, como se muestra en la encuesta del anexo 2.

- Las interfaces de manejo del sistema serán de fácil acceso e intuitivas para el usuario.
- El sistema web utiliza algoritmos de minería de datos, que brinden al visitante mayor información turística de cada atractivo que puede visitar.

## CAPÍTULO 2

# FUNDAMENTOS TEÓRICOS DE LA MINERÍA DE DATOS

## **2.1. CONCEPTOS Y DEFINICIONES DE LA MINERÍA DE DATOS**

La minería de datos o Data Mining es un término relativamente moderno que integra numerosas técnicas de análisis de datos y extracción de modelos. Son aplicables en varias disciplinas con el fin de extraer patrones, de describir tendencias y regularidades, de predecir comportamientos y, en general, de sacar partido a la información computarizada que nos rodea hoy en día, generalmente heterogénea. Con esto a su vez permite a los individuos y a las organizaciones comprender y modelar de una manera más eficiente y precisa el contexto en el que deben actuar y tomar decisiones.

Para definir podemos decir que “Data Mining es el conjunto de técnicas y herramientas aplicadas al proceso no trivial de extraer y presentar el conocimiento implícito, previamente desconocido, potencialmente útil y humanamente comprensible, a partir de grandes conjuntos de datos, con el objeto de predecir de forma automatizada tendencias y comportamientos y/o descubrir de forma automatizada modelos previamente desconocidos” [7].

## **2.2. TIPOS DE MODELOS DE MINERÍA DE DATOS**

En la práctica, los modelos pueden ser de dos tipos: predictivos y descriptivos.

### **2.2.1. MODELO PREDICTIVO**

Los modelos predictivos pretenden estimar valores futuros o desconocidos de variables de interés conocidas como variables objetivo o dependientes, usando otras variables o campos de la base de datos a las que llamamos variables independientes o predictivas [6].

### **2.2.2. MODELO DESCRIPTIVO**

Los modelos descriptivos en cambio, identifican patrones que explican o resumen los datos, es decir, sirven para explorar las propiedades de los datos examinados, no para predecir nuevos datos [6].

## 2.3. TAREAS Y TÉCNICAS DE LA MINERÍA DE DATOS

Dentro de la minería de datos se han de distinguir tipos de tareas, cada una de las cuales puede considerarse como un tipo de problema a ser resuelto a través de un método o técnica de minería de datos aplicando algoritmos. Esto significa que cada tarea tiene sus propios requisitos y que el tipo de información obtenida con una tarea puede diferir mucho de la obtenida con otra [6].

A continuación la tabla 2.1 muestra las distintas tareas existentes, las cuales pueden ser de dos tipos: predictivas o descriptivas, según el modelo que se desee obtener:

TAREA	DESCRIPCIÓN	TIPO
Clasificación	Es quizás la tarea más utilizada, consiste en encontrar una función que "mapea" (clasifica) un elemento de datos dentro de una de varias clases predefinidas.	Predictiva
Regresión	Consiste en aprender una función real que asigna a cada instancia un valor aproximado, minimizando el error cuadrático <sup>3</sup> .	Predictiva
Agrupamiento (clustering)	Es una tarea descriptiva por excelencia y consiste en obtener grupos "naturales" a partir de los datos.	Descriptiva
Correlación	Se usa para examinar el grado de similitud de dos variables numéricas.	Descriptiva
Reglas de asociación	Muy similar a las correlaciones, y tiene como objetivo identificar relaciones no explícitas entre atributos categóricos.	Descriptiva
Reglas de asociación secuenciales	Son un caso especial de reglas de asociación, que se usan para determinar patrones secuenciales en los datos.	Descriptiva

**Tabla 2.1.** Tareas de la minería de datos. [6]

---

<sup>3</sup> Se define como el cuadrado de la diferencia entre el valor real y el valor aproximado.

Cada una de las tareas anteriores requiere: métodos, algoritmos o técnicas para resolverlas. Una tarea puede ser resuelta por diferentes métodos y además tenemos que el mismo método puede resolver diferentes tareas. La tabla 2.2 muestra una clasificación de las técnicas de minería de datos más utilizadas en el desarrollo de aplicaciones.

TÉCNICA	DESCRIPCIÓN	ALGORITMO
Técnicas algebraicas y estadísticas	Expresan modelos y patrones mediante fórmulas algebraicas, funciones lineales, funciones no lineales, distribuciones o valores agregados	<ul style="list-style-type: none"> <li>• Regresión lineal</li> <li>• Regresión logarítmica</li> <li>• Regresión logística</li> </ul>
Técnicas bayesianas	Buscan determinar relaciones causales que expliquen un fenómeno según los datos contenidos en una base de datos. Se han usado principalmente para realizar predicciones.	<ul style="list-style-type: none"> <li>• Naive Bayes</li> <li>• Métodos basados en máxima verosimilitud</li> <li>• Algoritmo EM</li> </ul>
Conteos de frecuencia y tablas de contingencia	Estas técnicas se basan en contar la frecuencia en que dos o más sucesos se presentan conjuntamente.	<ul style="list-style-type: none"> <li>• Algoritmo a priori</li> </ul>
Árboles de decisión y sistemas de aprendizaje de reglas	Son estructuras que representan conjuntos de decisiones, y estas decisiones generan reglas para la clasificación de un conjunto de datos.	<ul style="list-style-type: none"> <li>• ID3</li> <li>• C4.5</li> <li>• CART</li> <li>• CN2</li> </ul>
Técnicas basadas en redes neuronales artificiales.	Se trata de técnicas que aprenden un modelo mediante el entrenamiento de los pesos que conectan un conjunto de nodos o neuronas.	<ul style="list-style-type: none"> <li>• Perceptrón Multicapa</li> <li>• Algoritmo de retropropagación</li> </ul>
Basadas en casos y en vecindad	Son métodos que se basan en distancias al resto de elementos, ya sea directamente como los vecinos más cercanos; o mediante la estimación de funciones de densidad.	<ul style="list-style-type: none"> <li>• K – Medias</li> <li>• Two-Step</li> </ul>

**Tabla 2.2.** Técnicas más usadas en la minería de datos. [6]

Como ejemplo se mencionan los algoritmos frecuentemente usados para resolver problemas de la minería de datos:

## Árboles de decisión ID3

El ID3 es un algoritmo simple y, sin embargo, potente, cuya misión es la elaboración de un árbol de decisión bajo las siguientes premisas:

- Cada nodo corresponde a un atributo y cada rama al valor posible de ese atributo. Una hoja del árbol especifica el valor esperado de la decisión de acuerdo con los ejemplos dados. La explicación de una determinada decisión viene dada por la trayectoria desde la raíz a la hoja representativa de esa decisión.
- A cada nodo es asociado aquel atributo más informativo que aún no haya sido considerado en la trayectoria desde la raíz.
- Para medir cuánto de informativo es un atributo se emplea el concepto de entropía<sup>4</sup>. Cuanto menor sea el valor de la entropía, menor será la incertidumbre y más útil será el atributo para la clasificación.

El ID3 es capaz de tratar con atributos cuyos valores sean discretos o continuos. En el primer caso, el árbol de decisión generado tendrá

---

<sup>4</sup> La entropía indica el límite teórico para la comprensión de los datos. También es una medida de la información contenida en el mensaje. La entropía o información transmitida por una distribución de probabilidad  $P = (p_1, p_2, \dots, p_n)$  viene dada por la siguiente función:  $E(P) = -(p_1 \log_2(p_1) + p_2 \log_2(p_2) + \dots + p_n \log_2(p_n))$ . Cuanto más uniforme sea la distribución de probabilidad, mayor será su información transmitida

tantas ramas como valores posibles tome el atributo. Si los valores del atributo son continuos, el ID3 no clasifica correctamente los ejemplos dados. [12].

### **Vecinos más próximos (K - mean)**

Es un método de agrupamiento por vecindad, basándose inicialmente en un número indeterminado de prototipos y de un conjunto de ejemplos a agrupar (ninguno de los elementos del conjunto tienen etiquetas o grupos previamente asignados). Los prototipos o centroides deben ser situados en el espacio, de tal forma que los datos pertenecientes al mismo centroide tengan características similares.

Para poder clasificar un nuevo ejemplo, previamente definidos los centroides correctamente, se compara el nuevo ejemplo con cada uno de los centroides y se asocia al que esté más próximo dependiendo del método de distancia elegida (la distancia más usada es la Euclídea<sup>5</sup>). [6]

---

<sup>5</sup> La distancia Euclídea es la distancia clásica como la longitud de la recta que une dos puntos en el espacio euclídeo:  $d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$  [6]



## Naive Bayes

Se basa en el teorema de Bayes<sup>6</sup>, el cual nos permite actualizar la creencia que tenemos de un suceso o conjunto de sucesos a partir de nuevos datos u observaciones. “Es decir, nos permite pasar de probabilidades a priori<sup>7</sup> a la probabilidad a posteriori<sup>8</sup>” [6]

Utiliza una estructura de red fija, de la cual es necesario aprender las probabilidades, los atributos que describen a los ejemplos son independientes entre sí con respecto al concepto que se pretende aprender (variable clase).

## Funciones Lineales Discriminantes

Estas consisten en dividir el espacio de estados, es decir, definir el conjunto de todos los posibles puntos en los que pueden ser ubicados los ejemplos en regiones, estas regiones están definidas por funciones (ecuaciones) lineales.

---

<sup>6</sup> Teorema de Bayes: es el resultado que da la distribución de probabilidad condicional de una variable aleatoria A dada B en términos de la distribución de probabilidad condicional de la variable B dada A y la distribución de probabilidad marginal de sólo A: 
$$P(A/B) = \frac{P(B/A)P(A)}{P(B)}$$

<sup>7</sup> A priori: (Del latín, ‘lo que viene antes de’), que se tiene información previa al proceso.

<sup>8</sup> A posteriori: (Del latín: ‘por lo que viene después’). Término que, a diferencia de a priori, designa el saber obtenido de la experiencia.

En la práctica, la asignación se realiza a partir de funciones monótonas de probabilidades a posteriori [6]

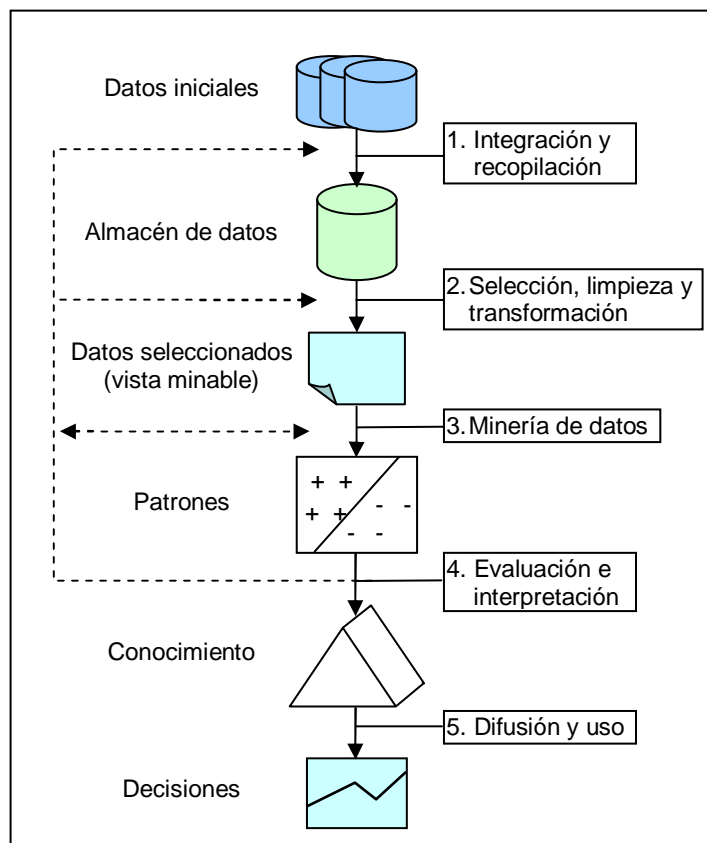
Además de las técnicas anteriores, existen multitud de híbridos que dificultan aún más realizar una clasificación acertada. En la tabla 2.3 se muestran algunas tareas y algunas técnicas o algoritmos que pueden abordarlas:

Nombre	PREDICTIVO		DESCRIPTIVO		
	Clasificación	Regresión	Agrupamiento	Reglas de Asociación	Correlaciones/ Factorizaciones
Redes Neuronales	✓	✓	✓		
Árboles de Decisión ID3, C4.5, C5.0	✓				
Árboles de decisión CART	✓	✓			
Otros árboles de decisión	✓	✓	✓	✓	
Redes de Coñeen			✓		
Regresión lineal y logarítmica		✓			✓
Regresión logística	✓			✓	
Kmeans			✓		
A priori				✓	
Naive Bayes	✓				
Vecinos más próximos	✓	✓	✓		
Análisis Factorial y de comp. princ.					✓
Twostep, Cobweb			✓		
Algoritmos genéticos y evolutivos	✓	✓	✓	✓	✓
Máquinas de vectores soporte	✓	✓	✓		
CN2 rules (cobertura)	✓			✓	
Análisis discriminante multivariante	✓				

**Tabla 2.3.** Tareas y tipo de técnica que pueden resolver dicha tarea. [6]

## 2.4. EL PROCESO DE EXTRACCIÓN DE CONOCIMIENTO

Es importante tener en cuenta que la minería de datos no es más que un paso esencial de un proceso más amplio cuyo objetivo es el descubrimiento de conocimiento en bases de datos. Este proceso cuenta con una secuencia de etapas o fases como lo muestra la fig. 2.2



**Figura 2.1.** Fases del proceso de extracción de conocimiento [6].

En la **fase de integración y recopilación de datos** se determinan las fuentes de información que pueden ser útiles y dónde conseguirlas. A continuación, se transforman todos los datos a un formato común, frecuentemente mediante un almacén de datos que consiga unificar de manera operativa toda la información recogida, detectando y resolviendo inconsistencias.

Dado que los datos provienen de diferentes fuentes, pueden contener valores erróneos o faltantes. Estas situaciones se tratan en la **fase de selección, limpieza y transformación**, en la que se eliminan o corrigen los datos incorrectos y se decide la estrategia a seguir con los datos incompletos. Además, se proyectan los datos para considerar únicamente aquellas variables o atributos que van a ser relevantes, con el objetivo de hacer más fácil la tarea propia de minería de datos y para que los resultados de la misma sean más útiles. Las dos primeras fases se suelen englobar bajo el nombre de “preparación de datos”.

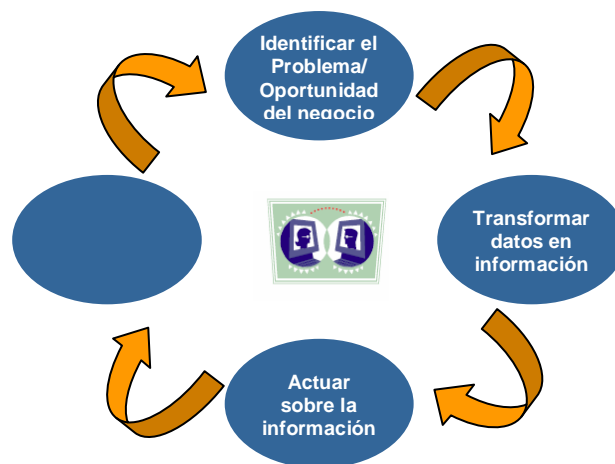
En la **fase de minería de datos** se decide cuál es la tarea a realizar (clasificar, agrupar, etc.) y se elige el método que se va a utilizar.

En la **fase de evaluación e interpretación** se evalúan los patrones y se analizan por los expertos, y si es necesario se vuelve a las fases anteriores para una nueva iteración.

Finalmente, en la **fase de difusión** se hace uso del nuevo conocimiento y se hace partícipe de él a todos los posibles usuarios.

## 2.5. EL CÍRCULO VIRTUOSO DE LA MINERÍA DE DATOS

El proceso de minería de datos puede verse como un ciclo virtuoso, basado en la continua mejoría de un proceso de negocios, impulsado por el descubrimiento de nuevo conocimiento y por la toma de acciones correspondientes a ese conocimiento [8].



**Figura 2.2.** Círculo virtuoso de la minería de datos

Los cuatro estados del ciclo virtuoso de la minería de datos [8], como se muestran en la figura 2.2, son:

**Identificar el problema/oportunidad del negocio.-** Aquí se analiza la temática del problema y se identifica donde los patrones en los datos tienen el potencial de proporcionar algún resultado.

**Transformar los datos en información.-** Consiste en usar técnicas de minería de datos para realizar su transformación en información procesable. Para este propósito, los resultados producidos necesitan ser interpretables para hacer que el ciclo virtuoso sea exitoso.

**Actuar sobre la información.-** Actuar sobre la información es el propósito del ciclo virtuoso de la minería de datos, es aquí donde los resultados de la minería de datos son usados para alimentar el siguiente paso.

**Medir los resultados.-** Finalmente se comparan los resultados obtenidos del modelo con los resultados deseados. Esta medida proporciona la retroalimentación para el mejoramiento continuo de los resultados y son las que hacen del círculo virtuoso de la minería de datos “virtuoso”.

Los puntos mencionados anteriormente serán analizados a medida que vayamos estudiando el problema en cada uno de los capítulos siguientes.

## **CAPÍTULO 3**

# **ANÁLISIS DEL SISTEMA**



### **3.1. ANÁLISIS Y REQUERIMIENTOS DEL SISTEMA**

En esta sección se detallan los requerimientos funcionales y de usabilidad que el sistema debe satisfacer.

El sistema proveerá información de interés acerca de lugares turísticos de la costa ecuatoriana, de manera tal que maximice la cobertura de exigencias en los usuarios y turistas en la búsqueda de sitios y atractivos.

Se espera realizar predicciones sobre los posibles atractivos que puede tomar el turista hacia aquellos lugares que den con su perfil turístico, así como también mostrar una categorización de los mejores sitios en base a una calificación dada por el turista que tenga conocimiento de los atractivos de la costa ecuatoriana.

Todas estas tareas necesarias en el sistema serán llevadas a cabo como un medio de información y de difusión de todo lo que acontece en cada uno de los parajes de la costa y como servicio pro-mejoras al turismo del país.

Para predecir el lugar adecuado se tenía pensado utilizar datos estadísticos con respecto al perfil de las personas que visitan los atractivos de la costa ecuatoriana, tales como: edad, sexo, estado civil, deportes favoritos, presupuesto, si viaja: solo, en pareja o en familia y las actividades que realizó durante su estadía.

Se realizaron encuestas a 75 personas de entre 18 y 50 años, con preguntas abiertas relacionadas a gustos y preferencias al momento de realizar turismo. Para verificar los resultados obtenidos, junto con el formato de la encuesta realizada, refiérase al anexo 2.

Estas encuestas ayudaron a encontrar un nuevo enfoque del problema, teniendo como nuevas variables las diferentes actividades que se realizan en cada atractivo y sus características como: lugar para comer, parqueo, concurrencia, limpieza y seguridad.

Las características y actividades del atractivo conforman las variables que se utilizan para generar el modelo de minería de datos propuesto.

Luego de haber analizado las necesidades de información y tomando en cuenta la disponibilidad de recursos con los que se cuenta, se ha planificado una aplicación de minería de datos: flexible, accesible,

escalable y confiable como una herramienta de: análisis, extracción, predicción, consulta, administración y distribución.

Esta herramienta ayudará a los turistas a encontrar la mejor ciudad que se acomode a su perfil turístico, de ahí nace la idea de crear un sistema de búsqueda en línea de sitios turísticos de la costa ecuatoriana.

En el sistema se ha considerado el desarrollo de los módulos a continuación descritos:

- *Búsqueda de Sitios Turísticos*: este módulo será capaz de: buscar ciudades con sus atractivos, mostrar atractivos turísticos y sitios turísticos por perfil, este último utilizará métodos de minería de datos que permitirán organizar y predecir los atractivos para determinado perfil del usuario.
- *Recomendación de Sitios Turísticos*: permite recomendar nuevos sitios y atractivos.
- *Calificación de Sitios Turísticos*: permite calificar un atractivo, teniendo un control de todas las calificaciones realizadas de cada uno de los atractivos previas para aplicar en el módulo de ranking la minería de datos.

- *Ranking de Sitios Turísticos*: permite tener un ranking de todos los sitios calificados, en base a la calificación del atractivo, su seguridad y limpieza.

### **3.1.1. ALCANCE DE DATOS**

El sistema abarcará dos (2) modelos de minería de datos, cumpliendo con los siguientes alcances:

El sistema permitirá al usuario ingresar las características del sitio requerido como factores para determinada búsqueda por perfil y aplicar la predicción como son:

- Lugar para comer
- Parqueo
- Seguridad
- Limpieza
- Concurrencia

Seleccionar las actividades que servirán para predecir el sitio más adecuado dependiendo de las preferencias del usuario y que se han clasificado como:

- Aventura
- Cultura y ciencia
- Deportes
- Deportes extremos
- Diversión nocturna
- Ecoturismo
- Gastronomía
- Recreación, y
- Otras actividades

Una vez realizada las operaciones necesarias, el usuario podrá observar la mejor predicción (ciudad) que se perfila a sus necesidades, junto con todos los atractivos que puede encontrar en el sitio. Además de otras predicciones (ciudades) más que se acercan a su perfil turístico.

Por ejemplo, si el usuario desea practicar algún deporte extremo y a la vez visitar un museo, le indicaremos cual es la ciudad más apropiada en la que puede quedarse para desde ahí visitar los atractivos turísticos en dónde podrá realizar las actividades seleccionadas.

Finalmente, el sistema proporcionará un ranking de todos los atractivos que hayan sido calificados por los usuarios, donde el usuario podrá informarse de los sitios que tengan mejor calificación.

### **3.1.2. ACCESO DE LOS CLIENTES AL SISTEMA**

El sistema de difusión de sitios turísticos tendrá dos tipos de usuarios los cuales tendrán diferentes privilegios:

- Administrador
- Usuario (turista o visitante)

El Administrador será el encargado de:

- Mantener la información actualizada periódicamente dentro del sistema.
- Agregar en la base de datos nuevos sitios, atractivos y actividades si así lo hubieren.
- Verificar que la calificación de los sitios contenga información coherente y adecuada para un mejor filtrado y para que, al realizar la minería, se produzcan los resultados esperados.

Mientras que los usuarios serán todos aquellos que deseen conocer sitios turísticos, sean estos turistas o visitantes, que buscan determinado lugar para vacacionar y realizar diferentes actividades turísticas dentro de un atractivo perteneciente a la costa ecuatoriana.

### **3.2. ANÁLISIS DE TÉCNICAS DE MINERÍA DE DATOS A IMPLEMENTARSE**

Al inicio de este capítulo se definieron las dos tareas de minería de datos que se desea resolver. La primera tarea es de tipo predictiva (predicción de sitio turístico), y la segunda tarea es de tipo descriptiva (ranking de sitios turísticos).

Como se explicó en el capítulo 2, existen muchas técnicas para resolver los problemas de minería de datos, de entre las cuales, las que mejor se acoplan al resolver tales tareas son:

- Árboles de decisión.
- Vecinos más próximos.
- Redes bayesianas (Naive Bayes).
- Funciones lineales discriminantes.

Fue necesario realizar varias pruebas previas con éstos algoritmos para determinar cuál era el más adecuado para el desarrollo del sistema. A continuación se describe como fue aplicado cada uno de éstos algoritmos para luego seleccionar el más adecuado según la tarea que se desea resolver.

### **Análisis de los algoritmos para la tarea predictiva**

Árboles de decisión (ID3), como se detalló en el capítulo 2, cada nodo corresponde a un atributo (actividad/tipo de atractivo) y cada *rama* al valor posible de ese atributo. Una *hoja* del árbol especifica el valor esperado de la predicción de acuerdo con los ejemplos dados. El algoritmo ID3 utiliza un recorrido podado, permitiendo escoger las variables (las actividades), con mayor entropía, utilizando elementos discretizados, para obtener un resultado.

El inconveniente con este método, aplicado a la tarea de predicción de sitio turístico, era que al realizar el recorrido del árbol, las variables escogidas por el usuario no necesariamente eran las variables que el algoritmo mantenía al entrenarse<sup>9</sup>, lo que produciría resultados

---

<sup>9</sup> Algoritmo de Árboles de Decisión con poda, hace una depuración de variables, eliminando las que considera que tienen menor información para el modelo.



incorrectos en la evaluación del modelo, como se demuestra en el anexo 10A.

Vecinos más próximos, se basa en teoremas de distancias para encontrar a que grupo pertenece un individuo (encajar el perfil del usuario, como se ha mencionado en el capítulo 2).

Este método es poco eficiente cuando se utilizan demasiadas variables, lo que significaba un gran problema para la implementación, ya que el usuario puede escoger cuantas variables (actividades) desee para la predicción, es decir, puede escoger desde una hasta todas las variables consideradas.

En el anexo 10B, se muestra un ejemplo, utilizando solamente 4 variables (“Bañarse en el mar”, “Visitar museos”, “Visitar ruinas arqueológicas” y “Ver aves”).

Predictor Naive Bayes, este método es el que mejor se ajusta a resolver el problema planteado, creando probabilidades de todas las variables (actividades y características del atractivo), las cuales al momento de ejecutar una predicción solo se seleccionan las de interés para el usuario, generando el resultado más apropiado.

Este método brinda una gran ventaja, dado que puede ser cambiado dependiendo de las variables que el usuario escoja para realizar la exploración al momento de ejecutarse, como se demuestra en el anexo 10C.

### **Análisis de los algoritmos para la tarea descriptiva.**

En el sistema de sitios turísticos de la costa ecuatoriana se aplica, como técnica, una función discriminante lineal para generar un modelo descriptivo que muestre el ranking de los mejores atractivos turísticos de la costa, utilizando variables como: la limpieza, la concurrencia y la calificación del atractivo, indicando los atractivos con mejor calificación según cada una de las variables antes mencionadas. Un ejemplo de la aplicación de este algoritmo se encuentra en el anexo 10D.

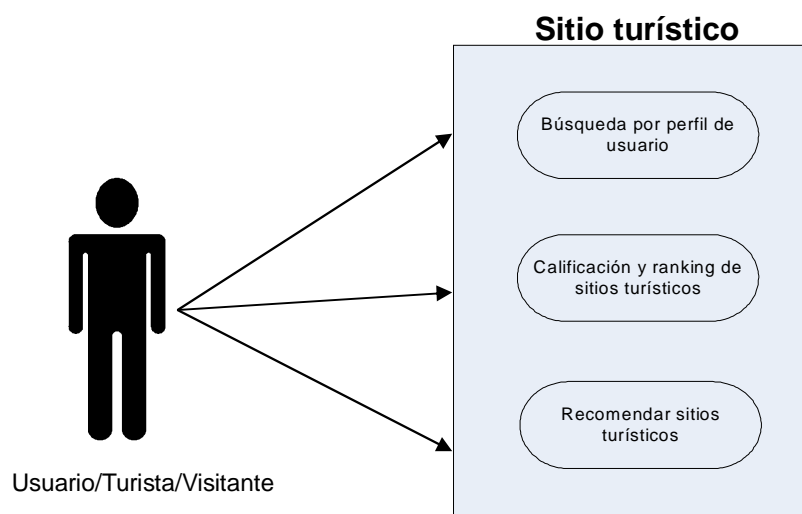
## **3.3. ANÁLISIS ORIENTADO A OBJETOS**

### **3.3.1. DIAGRAMA DE CASOS DE USO**

En el análisis del sistema de búsqueda en línea de sitios turísticos, se describe que hace el sistema desde el punto de vista del usuario. Para este sistema los casos de uso encontrados son los siguientes:

1. Búsqueda por perfil de usuario
2. Calificación y ranking de sitios turísticos
3. Recomendar sitios turísticos

Con estos casos de uso interactúa el actor Usuario/turista/visitante como se muestra en la siguiente figura:



**Figura 3.1.** Diagrama de casos de uso

### 3.3.2. ESPECIFICACIONES DE CASOS DE USO

A continuación, se detallan cada uno de los casos de uso del sistema:

<b>Caso de Uso 1:</b>	Búsqueda por perfil de usuario.
<b>Actor:</b>	Usuario (visitante o turista) del sistema.
<b>Descripción:</b> Este caso de uso describe el escenario donde el usuario digita la dirección web del sistema para conocer el sitio turístico y su funcionalidad.	
<b>Notas:</b> <ul style="list-style-type: none"> <li>• El usuario debe ingresar y seleccionar previamente algunos parámetros (características del sitio y actividades a realizar)</li> <li>• Si la búsqueda es exitosa se muestra al usuario él o los sitios de interés para éste.</li> <li>• Si la búsqueda no es exitosa no se presenta ningún sitio al usuario.</li> <li>• El sistema puede hacer tantas predicciones como el usuario lo requiera.</li> </ul>	

<b>Caso de Uso 2:</b>	Calificación y ranking de sitios turísticos.
<b>Actor:</b>	Usuario (visitante o turista) del sistema.
<b>Descripción:</b> En este caso de uso, el usuario realiza una votación o calificación de los sitios turísticos en base a ciertos parámetros a considerar. De esta forma, a la hora de pensar ver más información de los sitios turísticos mostrados tendrá como vista inicial la puntuación o ranking de los mejores sitios calificados.	
<b>Notas:</b> <ul style="list-style-type: none"> <li>• El usuario deberá llenar un formulario en el que realiza una calificación del sitio en base a ciertos parámetros.</li> <li>• El sistema recibe el formulario ingresado, valida los datos ingresados por el usuario y los actualiza en la base de datos</li> </ul>	

<b>Caso de Uso 3:</b>	Recomendar sitios turísticos.
<b>Actor:</b>	Usuario (visitante o turista) del sistema.
<b>Descripción:</b> Este caso de uso describe el escenario donde el usuario desea dar una opinión y recomendar un nuevo atractivo turístico, o una nueva ciudad.	
<b>Notas:</b> <ul style="list-style-type: none"> <li>• Para recomendar una ciudad el usuario deberá ingresar el nombre de la nueva ciudad creada, seleccionar la provincia en la que se encuentra, ingresar una descripción breve de la ciudad, cargar una imagen de la ciudad e ingresar un identificador que puede ser un nombre o nick que identifique al autor de la recomendación.</li> <li>• Para la recomendación de un nuevo atractivo será necesario seleccionar la ubicación del atractivo como: provincia y ciudad, e ingresar los datos del atractivo como: el nombre, una descripción, dirección, horario de atención, datos adicionales, insertar una imagen del atractivo, el identificador del autor de la recomendación y seleccionar además si existiera: lugar para comer, parqueo, el nivel de seguridad, nivel de limpieza y el nivel de concurrencia.</li> </ul>	

Para ver todos los escenarios de los casos de uso descritos anteriormente refiérase al Anexo 7.

### **3.4. ANÁLISIS DE LAS HERRAMIENTAS DE DESARROLLO**

#### **3.4.1. CRITERIOS PARA LA ELECCIÓN DE LAS HERRAMIENTAS**

Para este proyecto de tesis, era necesario escoger una tecnología actual, que soporte la manipulación de información necesaria de la base de sitios turísticos aplicando minería de datos y a la vez que permita crear una interfaz que sea agradable al usuario.

#### **3.4.2. JUSTIFICACIÓN PARA LA ELECCIÓN DE LAS HERRAMIENTAS SELECCIONADAS**

Para realizar este proyecto de tesis se decidió trabajar con la herramienta ASP.NET como aplicación y SQL Server como repositorio de datos, por que son una buena alternativa que brinda muchas facilidades además de su fácil aprendizaje y buen desempeño.

#### **Asp.Net**

Las siglas ASP corresponden a las palabras *Active Server Pages* (Páginas Activas en el Servidor), es un completo marco de trabajo de construcción de aplicaciones web. Una de las principales

características por las que fue escogida esta herramienta es la flexibilidad de escoger su lenguaje de programación.

Otra de las características más importantes de las páginas ASP es la posibilidad de conectar con diferentes tipos de bases de datos, para extraer, agregar o eliminar datos de ellas y generar páginas con esos datos. Se generan en forma dinámica dependiendo de las sentencias que se establezcan, para así obtener los resultados del proceso realizado.

Puede conectarse a motores de bases de datos SQL, Microsoft Access, Oracle, y a cualquier otro con soporte de conexión ODBC el cual proporciona un conjunto consistente de funciones para acceder a su base de datos.

Sus funcionalidades permiten un acceso mas sencillo y con la facilidad de poder emplear diferentes lenguajes de codificación pero sin perder potencia de diseño ya que los principales: C#, Visual Basic.NET y J#.NET son orientados a objetos. Además de que soportan herencias, polimorfismos y encapsulación. Su diseño e implementación es limpio y presenta un ambiente sofisticado de depuración de errores.

Dentro de las mejoras que presenta trabajar con ASP.net están:

- Rendimiento: la aplicación se compila en una sola vez al lenguaje nativo, y luego, en cada petición tiene una compilación Just In Time, es decir se compila desde el código nativo, lo que permite mucho mejor rendimiento. También permite el almacenamiento del caché en el servidor.
- Rapidez en programación: mediante diversos controles, podemos con unas pocas líneas mostrar toda una base de datos y hacer rutinas complejas.
- Servicios web: trae herramientas para compartir datos e información entre distintos sitios.
- Seguridad: tiene diversas herramientas que garantizan la seguridad de nuestras aplicaciones.

Como servidor de base de datos, se utilizó SQL Server 2000 debido a que soporta grandes cantidades de información, son sistemas bien seguros y la probabilidad de pérdida de información es muy poca.



## Microsoft SQL Server 2000

A continuación se mencionan algunas de las más importantes características del sistema de gestión de bases de datos ofrecido por Microsoft [10]:

- Seguridad: SQL Server 2000 presenta un conjunto de nuevas y sofisticadas características de seguridad: seguridad eficaz y flexible basada en funciones para servidores, bases de datos y perfiles de aplicaciones; herramientas integradas para la auditoría de la seguridad, con la posibilidad de realizar el seguimiento de 18 sucesos de seguridad distintos y subsucesos adicionales. SQL Server 2000 posee el certificado de nivel C2.<sup>10</sup>
- Alojamiento de aplicaciones: Gracias a la compatibilidad con varias instancias, SQL Server le permite aprovechar totalmente sus inversiones en hardware de forma que múltiples aplicaciones se pueden ejecutar en un solo servidor o externamente.

---

<sup>10</sup> El nivel C2 es una calificación dada por el gobierno de USA por la discreción de sus recursos protegidos y la capacidad de ser auditable. Es el nivel más alto disponible.

- Escalabilidad: SQL Server 2000 permite realizar un escalamiento hasta 32 CPU y 64 gigabytes (GB) de RAM, siendo capaz de manejar al máximo multiprocesamiento simétrico aprovechando al máximo el hardware
- Administración de bases de datos simplificada: Las características de ajuste y mantenimiento automáticos permiten a los administradores centrar su atención en otras tareas críticas.
- Alta Disponibilidad: SQL Server 2000 (64-bits) soporta hasta ocho nodos en un cluster, permitiendo configuraciones de alta disponibilidad más complejas. Al permitir que otros servidores de 64 bits puedan recuperar la ejecución de aplicaciones críticas en caso de paradas del servicio planificadas o no previstas, las operaciones de negocio están más protegidas, y las aplicaciones empresariales están disponibles todo el tiempo.

A pesar de que Asp.Net como SQL 2000 son tecnologías relativamente fáciles de aprender e implementar debido a la gran información proporcionada en manuales y a través de la web, es importante tener un conocimiento básico de él o los lenguajes de

programación que éstas soportan y en el caso de las bases de datos un conocimiento en el modelamiento de datos.

## **CAPÍTULO 4**

# **DISEÑO DEL SISTEMA**

## 4.1. VISIÓN GENERAL DEL DISEÑO

La figura 4.1 esquematiza la arquitectura del sistema en el cual se muestra la relación entre el cliente, el servidor web y la base de datos:

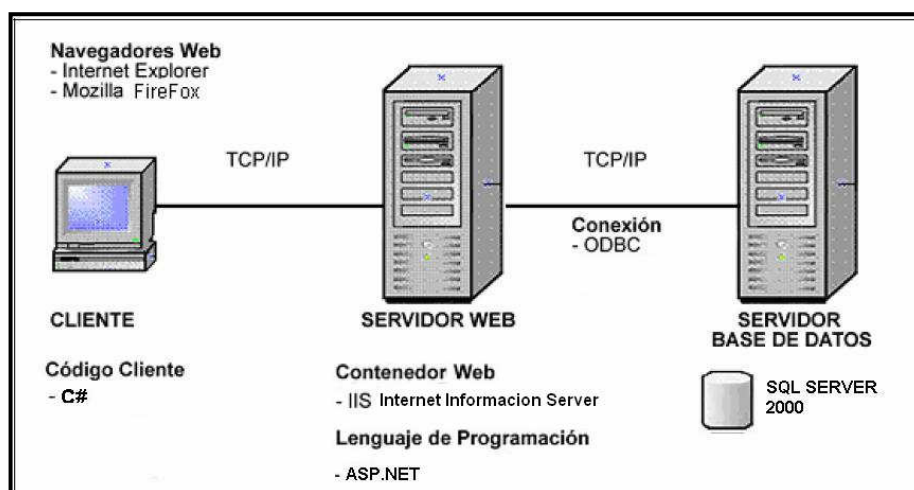


Figura 4.1. Arquitectura del Sistema

### Navegador cliente

Muestra las páginas Aspx al cliente con la información contenida en ellas para la navegación del sistema y manipular el contenido que está presente de manera dinámica e interactiva.

### Servidor Web

Es el encargado de atender los pedidos de los navegadores clientes. Realizará las validaciones necesarias de la información que el cliente

ingrese en los formularios, y devolverá respuesta al usuario. En este servidor se utilizó Internet Information Server (IIS).

Como lenguaje de programación para implementar los objetos, clases y estructuras del sistema en conjunto con las páginas, se utilizó Asp.Net con C#.

### **Servidor de base de datos**

En él se guarda todo el repositorio de datos sobre el cual se aplicará las técnicas de minería de datos para extraer, importar y analizar los datos.

Como herramienta de base de datos se utilizó Microsoft SQL 2000.

## **4.2. DISEÑO ORIENTADO A OBJETOS**

Una vez realizado el análisis del sistema, con sus diagramas y especificaciones de casos de uso, descritos en el capítulo 3, pasamos al desarrollo del diseño del sistema, para el cual se ha considerado un diagrama de secuencias por cada escenario en el que se determinan las clases y asociaciones que se utilizarán en la implementación, así como los algoritmos y métodos utilizados para implementar las operaciones.

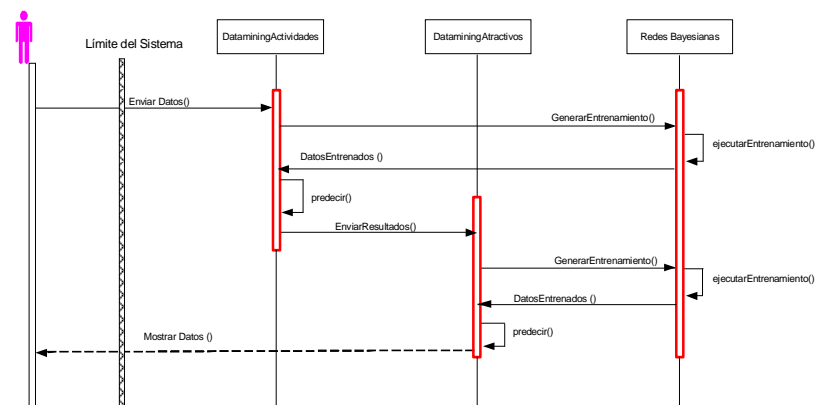
En el anexo 6 puede encontrar el diagrama de clases del sistema.

### 4.2.1. DIAGRAMAS DE INTERACCIÓN DE OBJETOS

A continuación se muestran cada uno de los escenarios posibles asociados a un determinado caso de uso del sistema.

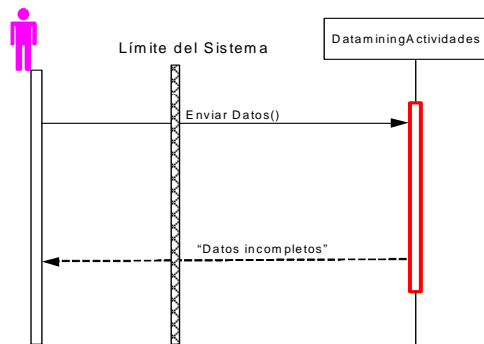
#### Caso de uso 1: Búsqueda por perfil de usuario.

Escenario 1.1: Sistema realiza búsqueda de sitios turísticos exitosamente



**Figura 4.2.** Sistema realiza búsqueda de sitios turísticos exitosamente

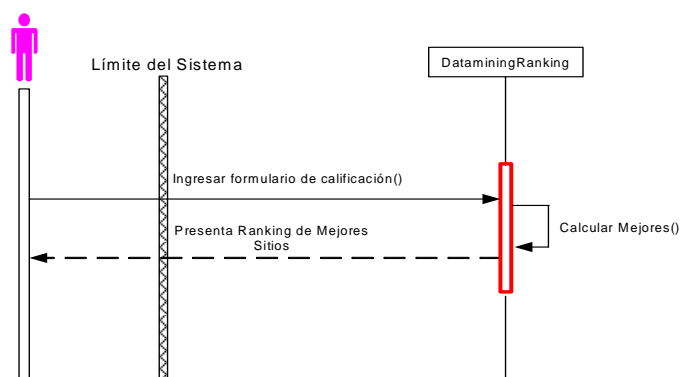
Escenario 1.2: Sistema no puede realizar búsqueda de sitios turísticos exitosamente.



**Figura 4.3.** Sistema no puede realizar búsqueda de sitios turísticos exitosamente

## Caso de uso 2: Calificación y ranking de sitios turísticos

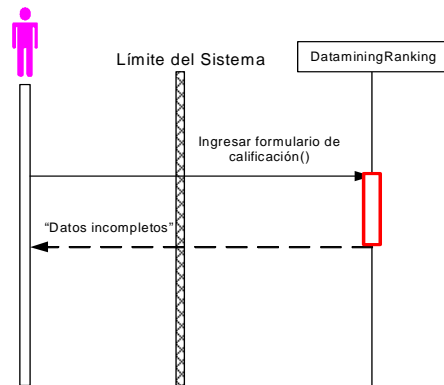
Escenario 2.1: Sistema actualiza calificación y ranking de sitios turísticos exitosamente



**Figura 4.4.** Sistema actualiza calificación y ranking de sitios turísticos exitosamente.



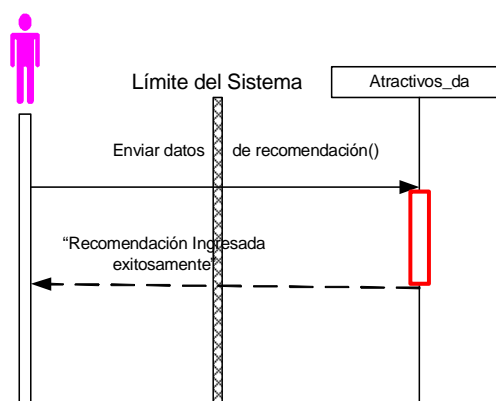
Escenario 2.2: Sistema no actualiza calificación y ranking de sitios turísticos exitosamente.



**Figura 4.5.** Sistema no actualiza calificación y ranking de sitios turísticos exitosamente.

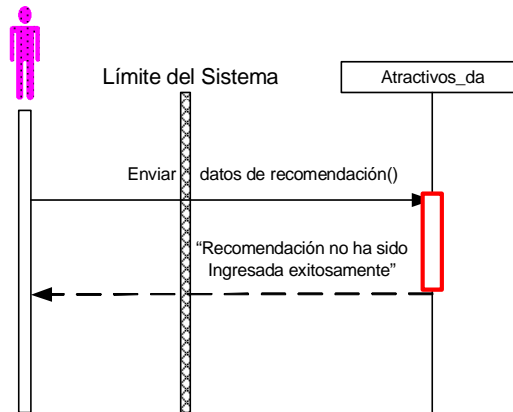
### Caso de uso 3: Recomendar sitios turísticos.

Escenario 3.1: Usuario ingresa recomendación de sitios turísticos exitosamente



**Figura 4.6.** Recomendar sitios turísticos exitosamente

Escenario 3.2: Usuario no puede ingresar recomendación de sitios turísticos exitosamente



**Figura 4.7.** No se puede recomendar sitios turísticos exitosamente

### 4.3. DISEÑO DE LA BASE DE DATOS

Los procesos y consultas que se realizan en el sistema, dependen de un buen modelamiento de la base de datos, ya que en ella se almacenará tanto la información, como la descripción de cada uno de los elementos que son parte del sistema.

#### 4.3.1. MODELAMIENTO DE LA BASE DE DATOS

A continuación en la tabla 4.1 se describen cada una de las entidades que intervienen en el modelo del sistema, junto como sus atributos.

Entidad	Atributos
st_provincia	idProvincia, nombre, descripción
st_ciudad	idCiudad, nombre, descripción, imagen, status
st_atractivo	Inatractivo, nombre, descripción, dirección, imagen, lugarComer, parqueo, horarioAtención, seguridad, limpio, concurrencia, estatus, datosAdicionales
st_tipoAtractivo	idTipoAtractivo, nombre, descripción
st_actividad	idactividad, nombre, descripción
st_categoria	idCategoría, nombre, descripción
st_revisión	idRevisión, descripción, calificación, pros, cont, limpieza, seguridad, posteador, fechaCreación

**Tabla 4.1** Entidades descriptivas del Modelo de la Base de Datos

La descripción de de cada uno de los atributos se encuentra detallada en el Anexo 9.

En cuanto a las relaciones entre las entidades y su cardinalidad, se describen como sigue en la tabla 4.2.

Relación	Entidades Participantes	Cardinalidad	Atributos de Relación
Ciudad <u>pertenece a</u> una Provincia	st_provincia st_ciudad	Uno a Muchos	No
Atractivo <u>pertenece a</u> una ciudad	st_ciudad st_atractivo	Uno a Muchos	No
Atractivo <u>pertenece a</u> un TipoAtractivo	st_tipoAtractivo st_atractivo	Uno a Muchos	No
Revisión <u>pertenece a</u> un Atractivo	st_atractivo st_revisión	Uno a Muchos	No
Actividad que <u>pertenece a</u> una Categoría	st_categoria st_actividades	Uno a Muchos	No
Actividades <u>se encuentran en</u> Atractivos	st_atractivo st_actividad	Muchos a Muchos	No

**Tabla 4.2** Relaciones entre entidades

Las claves primarias van a ser un atributo secuencial que va a servir para identificar de manera única a los registros:

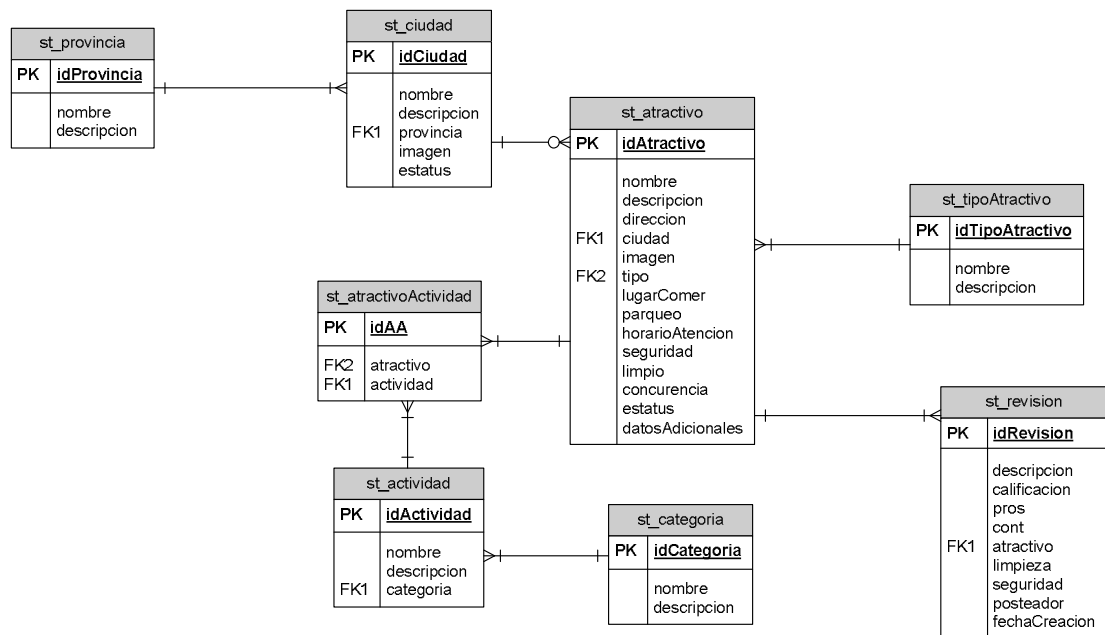
- st\_provincia: idProvincia
- st\_ciudad: idCiudad
- st\_atractivo: idAtractivo
- st\_tipoAtractivo: idTipoAtractivo
- st\_actividad: idActividad
- st\_categoria: idCategoría
- st\_revisión: idRevisión

La entidad intermedia para describir las relaciones de muchos a muchos es la siguiente tabla:

Relación	Entidad Intermedia	Atributos
Actividades se <u>encuentran en</u> Atractivos	st_atractivoActividad	idAA, atractivo, actividad

**Tabla 4.3** Relación entre entidades intermedias

El diagrama final de la estructura de la base de datos de nuestro sistema o diagrama entidad-relación se muestra en la figura 4.8



**Figura 4.8.** Diseño de la base de datos sitios turísticos

#### 4.3.2. DISEÑO MULTIDIMENSIONAL DE LA BASE DE DATOS

En esta sección se describe el diseño multidimensional propuesto para realizar la minería de datos. A continuación se detallan las tablas de hechos y de dimensión involucradas en este modelo.

Las tablas de hechos de la multidimensionalidad son: md\_atractivoCalificación y md\_ciudadCalificación, las mismas que se describen en la tabla 4.4.

Entidad	Atributos
md_ciudadCalificación	<u>Id</u> , calificaciónSeguridad, calificaciónLimpieza, calificaciónCiudad, puntajeCampo, puntajeCultura, puntajeMonumentos, puntajeComercio, puntajeReservas, ranking, parqueoProb, lugarComidaProb, concurrenciaProb
md_atractivoCalificación	<u>Id</u> , puntajeDeportes, puntajeDepEx, puntajeDivNoct, puntajeCultura, puntajeEcoturismo, puntajeAventura, puntajeRecreación, calificaciónSeg, calificaciónConcurrencia, calificaciónAtractivo

**Tabla 4.4** Tablas de Hecho

Las tablas de dimensión utilizadas en el sistema se detallan en la tabla 4.5.

Entidad	Atributos
md_semana	<u>Id</u> , semana, fecha
md_mes	<u>Id</u> , mes
md_año	<u>Id</u> , año
md_categoria	<u>idCategoría</u> , nombre, descripción
md_actividad	<u>idActividad</u> , nombre, descripción
md_tipoAtractivo	<u>idTipoAtractivo</u> , nombre, descripción
md_atractivo	Id, nombre, dirección, lugarComer, parqueo, seguridad, limpio, concurrencia
md_ciudad	<u>idCiudad</u> , nombre, provincia
md_provincia	<u>idProvincia</u> , Nombre y descripción

**Tabla 4.5** Tablas de dimensión

En cuanto a las relaciones entre las entidades y su cardinalidad, se describen como sigue en la tabla 4.2.

Relación	Entidades Participantes	Cardinalidad	Atributos de Relación
Ciudad <u>pertenece a</u> una Provincia	md_provincia md_ciudad	Uno a Muchos	No
Atractivo <u>pertenece a</u> una ciudad	md_ciudad md_atractivo	Uno a Muchos	No
Atractivo <u>pertenece a</u> un TipoAtractivo	md_tipoAtractivo md_atractivo	Uno a Muchos	No
Actividad que <u>pertenece a</u> una Categoría	md_categoría md_actividades	Uno a Muchos	No
Actividades <u>se encuentran en</u> Atractivos	md_atractivo md_actividad	Muchos a Muchos	No
Semana <u>pertenece a</u> Mes	md_semana md_mes	Uno a Muchos	No
Mes <u>pertenece a</u> Año	md_mes md_año	Uno a Muchos	No
Semana <u>se encuentra en</u> md_ciudadCalificación	md_semana md_ciudadCalificación	Uno a Muchos	No
Semana <u>se encuentra en</u> md_atractivoCalificación	md_semana md_atractivoCalificación	Uno a Muchos	No
md_ciudadCalificación <u>tiene una</u> ciudad	md_ciudadCalificación md_ciudad	Muchos a Uno	No
md_atractivoCalificación <u>tiene un</u> atractivo	md_atractivoCalificación md_atractivo	Muchos a Uno	No

**Tabla 4.6** Relaciones entre entidades del modelo multidimensional

Las claves primarias para este modelo, van a ser un atributo secuencial que va a servir para identificar de manera única a los registros:

- md\_provincia: idProvincia
- md\_ciudad: idCiudad
- md\_atractivo: id
- md\_tipoAtractivo: idTipoAtractivo
- md\_actividad: idActividad

- md\_categoria: idCategoría
- md\_semana: id
- md\_mes: id
- md\_año: id
- md\_atractivoCalificacion: id
- md\_ciudadCalificacion: id

La entidad intermedia para describir las relaciones de muchos a muchos se muestra en la tabla 4.7:

Relación	Entidad Intermedia	Atributos
Actividades <u>se encuentran en</u> Atractivos	md_atractivoActividad	idAA, atractivo, actividad

**Tabla 4.7** Relación entre entidades intermedias

La figura 4.9 muestra el diagrama del diseño multidimensional de la base de datos propuesto para el sistema. En este diseño se aprecian las tablas de hechos: md\_atractivoCalificación y md\_ciudadCalificación.



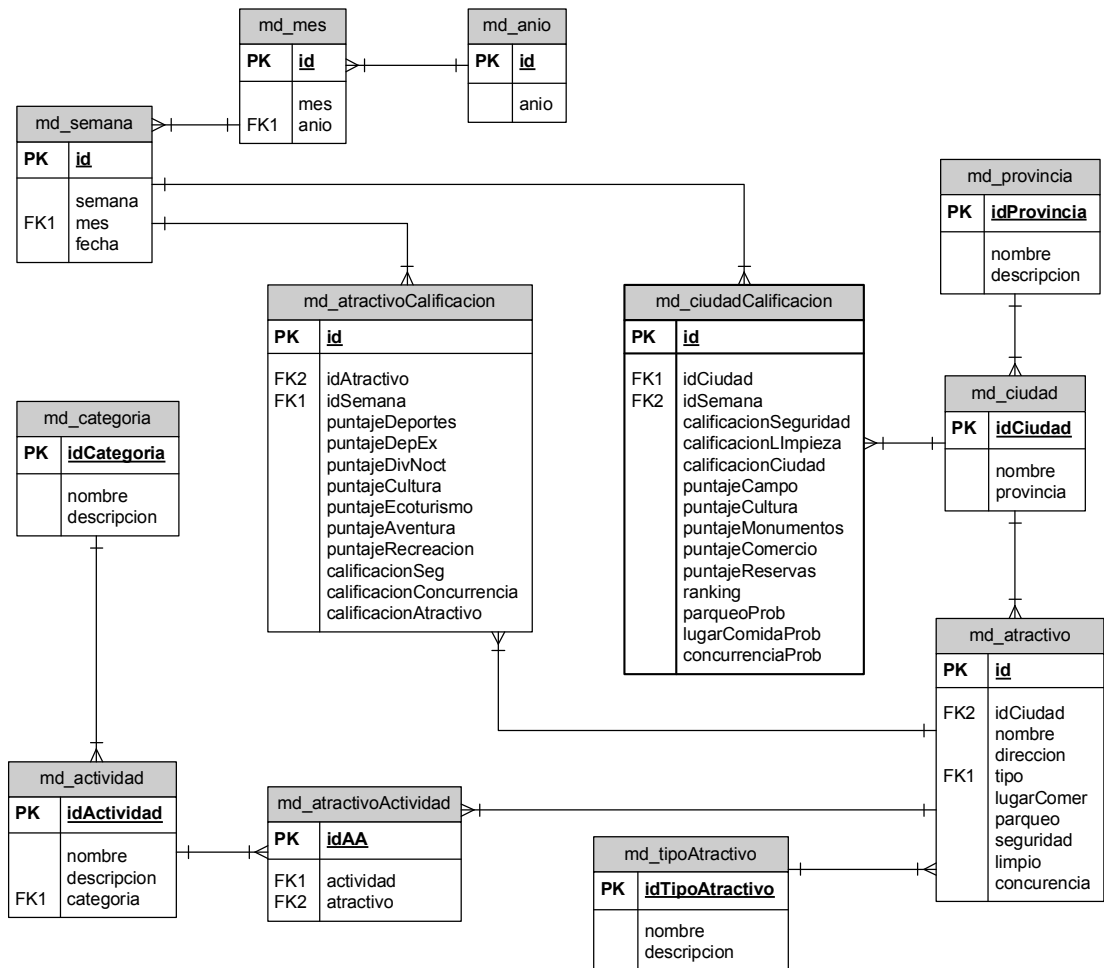


Figura 4.9. Diseño multidimensional de la base de datos

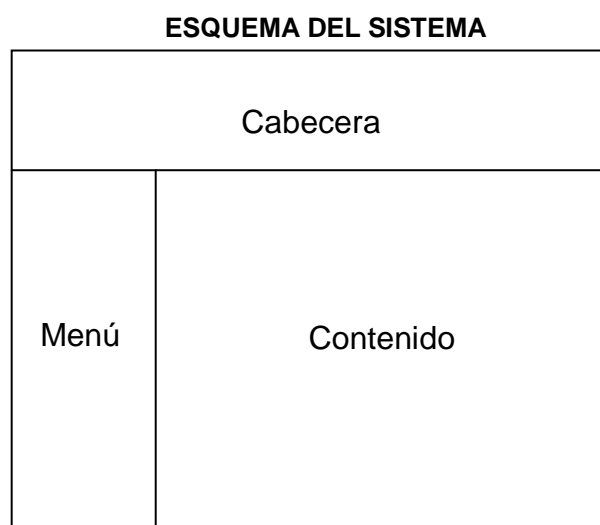
#### 4.4. DISEÑO DE LA APLICACIÓN Y LA INTERACCIÓN CON EL USUARIO

En esta sección se define el diseño de la interfaz con el usuario, considerando algunos factores de usabilidad como: la visibilidad, la familiaridad, la consistencia, la flexibilidad y la satisfacción del usuario.

El diseño del sistema de búsqueda en línea de sitios turísticos se lo realizó no solo pensando en las búsquedas, sino también, se añadieron algunas opciones como son: el ranking, la calificación de sitios turísticos, y la recomendación de nuevas ciudades y atractivos de la costa ecuatoriana.

Las páginas que conforman el sistema están organizadas como se muestra en el esquema de la figura 4.10.

En la parte superior se encuentra la cabecera que contiene el logotipo y nombre del sistema y en la parte inferior izquierda de la cabecera se muestra al usuario la fecha actual; al lado izquierdo, debajo de la cabecera, está ubicado el menú que despliega las opciones que presenta el sistema, siguiendo los estándares de diseños de páginas web; finalmente, en la zona central de la página se presentan el contenido de acuerdo a la opción que escoja el usuario.



**Figura 4.10.** Esquema del diseño del sistema

El sistema fue hecho de tal manera que al usuario le sea fácil recordar donde se encuentra cada opción, además los menús que se despliegan de cada botón siguen la misma secuencia, lo que hace que su uso sea mucho más fácil para el usuario. Otra consideración que se tomó fue la de mostrar un solo menú cada vez que se presiona la cabecera del menú, dando así al usuario el enfoque a la opción escogida.

Los colores para la aplicación, fueron seleccionados usando una paleta basada en tonos azules, tomando en cuenta que dicho color identifica la costa ecuatoriana bañada por el océano; de tal manera que el usuario se sienta a gusto al usar el sistema. El tipo de letra escogido para el

sistema es totalmente legible, de modo que el usuario no tenga problemas en la comprensión de los datos e información mostrados.

Para indicar al usuario los errores en las acciones como en el ingreso y edición de datos en formularios, se muestran mensajes descriptivos ubicados en la parte superior del formulario.

Para visualizar de forma general el flujo de ventanas del sistema puede revisar el anexo 5.

## **CAPÍTULO 5**

# **IMPLEMENTACIÓN Y PRUEBAS DEL SISTEMA**

## **5.1. IMPLEMENTACIÓN DE MÉTODOS DE MINERÍA DE DATOS**

En el desarrollo de esta tesis se implementaron dos técnicas de minería de datos para resolver los problemas planteados. En esta sección se detallan su desarrollo y los resultados obtenidos.

### **5.1.1. IMPLEMENTACIÓN DE TÉCNICAS BAYESIANAS PARA EL MODELO PREDICTIVO**

Para el diseño del modelo predictivo se decidió implementar técnicas bayesianas como el algoritmo Naive Bayes, ya que se conoce las diferentes clases o grupos donde se puede realizar la predicción.

En la predicción se consideraron dos fases, y en cada una de ellas se aplicó Naive Bayes. La primera fase abarca perfilar al usuario de acuerdo a las actividades que escoja, y como segunda fase la predicción de la ciudad que mejor encaje con las actividades seleccionadas en la primera fase más las variables del atractivo como son: lugar para comer, parqueo, seguridad, limpieza y concurrencia.

### **Primera fase – predicción del tipo de atractivo**

En la primera fase se perfila al usuario, en base a las actividades que ingresa (deportes, deportes extremos, aventura, gastronomía, etc.), para aproximar el tipo de atractivo que está buscando.

Se han dividido los tipos de atractivos turísticos en seis categorías: campo, balneario, cultural, parques y malecones, comerciales y reservas ecológicas. Esto indica que un atractivo puede entrar en varias categorías dependiendo de las diferentes actividades que se puedan realizar dentro de éste.

La variable clase dada que interviene en la primera fase del sistema es el ***tipo de atractivo***, que conjuntamente con las variables (actividades) que escoja el usuario, se procesan y se filtran los datos de nuestro repositorio, y se usan los datos óptimos para alimentar el modelo. Es así que como resultado de aplicar Naive Bayes se estima y predice los tipos de atractivos.

Para la implementación del algoritmo de Naive Bayes en la primera fase se supone tomar atributos independientes como variables, en este caso utilizamos las actividades que se realizan en un atractivo

(ver anexo 3), y como variable clase, el tipo de atractivo (ver anexo 4). Estos datos se ingresan discretizados, es decir, como una matriz de ceros (si es que no se realiza esa actividad en el atractivo) o 1 (si se realiza la actividad en el atractivo) como se muestra en la tabla 4.1. Para revisar la tabla completa de las variables utilizadas revisar el anexo 12.

		VARIABLES				VARIABLE CLASE
		Observar vuelo de gaviotas (V1)	Visitar museos (V2)	...	Admirar fauna (Vn)	TIPO ATRACTIVO (C)
ATRATIVOS		0	1		1	1
		0	1		0	1
		1	1		1	1
		0	1		0	1
		<b>Total incidencias =1</b>	<b>Total incidencias =4</b>		<b>Total incidencias =2</b>	
	<b>P(V1/C=1)=0,25</b>	<b>P(V2/C=1)=1</b>	<b>...</b>	<b>P(Vn/C=1)=0,5</b>		
ATRATIVOS						
		1	1		1	2
		0	1		0	2
		0	0		0	2
		<b>Total incidencias =1</b>	<b>Total incidencias =2</b>		<b>Total incidencias =1</b>	
	<b>P(V1/C=2)=0,33</b>	<b>P(V2/C=2)=0,67</b>	<b>...</b>	<b>P(Vn/C=2)=0,33</b>		
ATRATIVOS		.	.	.	.	.
		.	.	.	.	.
		.	.	.	.	.
	<b>P(V1/C=k)</b>	<b>P(V2/C=k)</b>	<b>...</b>	<b>P(Vn/C=k)</b>		

**Tabla 5.1.** Matriz de registros



Con estos datos se procede a crear la matriz de probabilidades (entrenamiento) del algoritmo bayesiano, la cual genera un vector de probabilidades por cada variable (actividad) asociados a los tipos de atractivos, como se muestra en la tabla 4.2.

Estos vectores son luego utilizados para la predicción del tipo de atractivo, tomando aquellos con más alta probabilidad según las actividades escogidas.

Vectores de Probabilidades
$V1 = \langle 0,25; 0,33; \dots P(V1/C=k) \rangle$
$V2 = \langle 1; 0,67; \dots P(V2/C=k) \rangle$
...
$Vn = \langle 0,5; 0,33; \dots P(Vn/C=k) \rangle$

**Tabla 5.2.** Vectores de probabilidades de cada variable

### **Segunda fase – predicción de mejor ciudad**

En la segunda fase del sistema, se predice la ciudad más apropiada para el usuario, utilizando el tipo de atractivo estimado en la fase anterior, conjuntamente con las características del atractivo: lugar para comer, parqueo, seguridad, limpieza y concurrencia; generando así una nueva matriz de probabilidades (entrenamiento).

Para la predicción de la mejor ciudad, se utilizan las características de los atractivos (lugar para comer, parqueo, seguridad, limpieza y concurrencia) como variables para generar el modelo, teniendo como variable clase, la **ciudad** a la que pertenece el tipo de atractivo que se predijo en el modelo anterior, todas del tipo discretas (contienen datos numéricos tipo entero), como se muestra en la tabla 4.3.

		VARIABLES				VARIABLE CLASE
		Lugar para comer (V1)	Parqueo (V2)	...	Limpieza (Vn)	CIUDAD (C)
ATRATIVOS		0	1		2	1
		0	1		3	1
		1	1		1	1
		0	1		0	1
		Total incidencias =1	Total incidencias =4		Total incidencias =1	
	$P(V1=1/C=1)=0,25$	$P(V2=1/C=1)=1$	...	$P(Vn=1/C=1)=0,25$		
ATRATIVOS		1	1		1	2
		0	1		3	2
		0	0		0	2
		Total incidencias =1	Total incidencias =2		Total incidencias =1	
	$P(V1=1/C=2)=0,33$	$P(V2=1/C=2)=0,67$	...	$P(Vn=1/C=2)=0,33$		
ATRATIVOS		.	.	.	.	.
		.	.	.	.	.
		.	.	.	.	.
	$P(V1=a/C=k)$	$P(V2=a/C=k)$	...	$P(Vn=a/C=k)$		

**Tabla 5.3.** Matriz de registros

Una vez obtenidos todos los datos del repositorio, se procede a la creación de las matrices de probabilidades, generando un vector de probabilidades por cada una de las características de los atractivos, como se muestra en la tabla 4.4, asociado a las ciudades, para determinar cuál es la ciudad mas apropiada para el turista.

Vectores de Probabilidades
<b>V1</b> = <0,25; 0,33;...P(V1/C=k)>
<b>V2</b> = <1; 0,67;... P(V2/C=k)>
...
<b>Vn</b> = <0,25; 0,33;... P(Vn/C=k)>

**Tabla 5.4.** Vectores de probabilidades de cada variable

Lo que se pretende con el algoritmo de Naive Bayes es construir una red de matrices con las probabilidades de ocurrencia de los sucesos, donde se da tratamiento especial a la variable clase y mejorar así la tasa de acierto durante la predicción.

### 5.1.2. IMPLEMENTACIÓN DE LA FUNCIÓN LINEAL DISCRIMINANTE PARA EL MODELO DESCRIPTIVO

Para la implementación del módulo de ranking del sistema se decidió utilizar una función lineal discriminante con información a priori, ya que permite clasificar los elementos en varios grupos (concurrentia,

limpieza y calificación del atractivo), según sea necesario, para luego obtener el mejor de cada grupo, es decir; se realiza una clasificación de los atractivos según sus calificaciones.

Inicialmente se realizan las siguientes asunciones:

- $K \leq n$
- Se tiene una muestra de tamaño N (número de registros de los atractivos con calificaciones)

Luego se agrupan las calificaciones de cada atractivo y se obtiene el promedio de calificaciones de cada uno, para aplicar la siguiente función discriminante:

$$y_k = \alpha_1 x_{1,i} + \alpha_2 x_{2,i} + \dots + \alpha_k x_{k,i} + \dots + \alpha_n x_{n,i}$$

Donde:

- n = cantidad total de atributos del elemento *i-ésimo*
- k = categoría que se está evaluando (conurrencia, limpieza y calificación del atractivo)

- $i$  = elemento *i-ésimo* que se está evaluando y que corresponde al promedio de calificación de cada atractivo (atractivo *i-ésimo* que se está evaluando)
- $\alpha$  = coeficiente de probabilidades.
- $N$  = Total de atractivos que han sido calificados.
- $X_{ki}$  = Calificación en la clase  $k$  del atractivo  $i$ .
- $K$  = número de categorías.

		Atractivos				
		1	2	I	...	N
Clase	1...	1	3	2	...	2
	K...	1	2	$X_{ki}$	...	1
	N	...	...	...	...	$X_{Nn}$

**Tabla 5.5.** Matriz de calificaciones de los atributos.

	Coeficientes de probabilidad		
	$\alpha_1$	$\alpha_2$	$\alpha_3$
<b>Calf. Atract.</b>	0,25	0,5	0,25
<b>Concurrencia</b>	0,25	0,25	0,5
<b>Limpieza</b>	0,5	0,25	0,25

**Tabla 5.6.** Coeficientes de probabilidad usados.

El factor  $\alpha$  toma un valor máximo dada la siguiente premisa:

$$\alpha_k > \alpha_1 + \alpha_2 + \dots + \alpha_{k-1} + \alpha_{k+1} + \dots + \alpha_n$$

Se efectúa la evaluación de las funciones  $Y_k \{y_1, y_2, \dots, y_k, \dots, y_n\}$  para el elemento *i-ésimo*, verificando cuál es la calificación más alta. En base a esto, se determina a que clase pertenece dicho elemento, usando:

$$y_{MAX} = MAX\{y_1, y_2, \dots, y_k, \dots, y_n\}$$

Donde el índice de  $y_{MAX}$ , indica la clase a la que pertenece el elemento *i-ésimo*. Una vez realizado este procedimiento se agrupan todos los  $x_i$  (atractivos) pertenecientes a cada clase, seleccionando el mayor para determinar el mejor de cada clase.

## 5.2. IMPLEMENTACIÓN DEL SISTEMA

Como se definió en el capítulo 4, las herramientas a utilizar para la implementación del sistema son: Visual Studio .NET, con los lenguajes de programación: ASP.NET y C#; debido a que se eligieron herramientas Microsoft para el desarrollo, se requirió un servidor Microsoft para la aplicación web, optando por el sistema operativo Microsoft Windows 2000 Server con el IIS; y, para el almacenamiento de la base de datos, se utiliza Microsoft SQL Server 2000.

Otras tecnologías utilizadas en la implementación son los javascripts (los cuales son usados del lado del cliente).

Para la implementación del sistema se utilizó un modelo basado en tres capas: capa de interfaz de usuario (GUI), capa de negocio y la capa de acceso a datos, las cuales fueron incorporadas en un mismo servidor.

Capa de interfaz de usuario: en esta capa se encuentran las páginas web, implementadas en ASP.NET, que proveen la interacción del usuario con el sistema y las validaciones hacia el usuario. Esta capa se comunica únicamente con la capa de negocio.

Capa de negocios: contiene la lógica del negocio, provista por los objetos que interactúan para realizar las tareas del sistema, esta capa esta implementada en C#, es aquí donde se establecen todas las reglas que deben cumplirse. Esta capa se comunica con la capa de interfaz de usuario, para recibir las solicitudes y presentar los resultados, y con la capa de acceso a datos, para solicitar al gestor de base de datos que almacene o recupere datos de él.

Capa de acceso a datos: esta capa provee la comunicación entre la base de datos y la aplicación, y es donde residen los datos. Está

formada por un motor de bases de datos que realiza todo el almacenamiento de datos, recibiendo solicitudes de almacenamiento o recuperación de información desde la capa de negocio.

Cada una de estas capas han sido aplicadas a la implementación del sistema se considerando los siguientes módulos: módulo de búsqueda de sitios turísticos, módulo de ranking de sitio turístico, módulo de recomendar nuevo sitio turístico y módulo de calificación de sitios

### **Módulo de búsqueda de sitios turísticos.**

Este módulo consta de tres partes: búsqueda de atractivo, búsqueda de ciudad y búsqueda por preferencias del turista.

Búsqueda de atractivo: una vez recolectada la información en el formulario de búsqueda, se llama a la función `buscarAtractivo(matrizRegistros11)`, la cual realiza la búsqueda de la información solicitada por el usuario.

Búsqueda de ciudad: para esta parte se utiliza la función `buscarCiudades(matrizRegistros)`, la cual realiza la búsqueda de la

---

<sup>11</sup> MatrizRegistro: es un objeto que representa lógicamente una estructura tipo matriz.



información solicitada por el usuario, una vez que haya sido recolectada la información en el formulario de búsqueda.

*Búsqueda por preferencias del turista:* en esta sección se realiza una predicción del lugar turístico que se acomode a las preferencias del turista. Una vez que el usuario ha seleccionado las características y actividades requeridas, éstos datos son usados por la función `dataMining()`, la cual realiza las dos fases de la predicción.

En la primera fase interviene el objeto `dataMinigActividades()`, que genera el entrenamiento para Naive Bayes y ejecuta la función `predecir(registro12)`, la cual recibe las variables(actividades) seleccionadas por el usuario y retorna la predicción del tipo de atractivo.

Para la segunda fase se toma el valor que retorna la función `predecir(registro)` de la fase anterior, y se llama al objeto `dataMinigAtractivo()`, que posee el entrenamiento para Naive Bayes y llama a su propia función `predecir(registro)`, la cual recibe las variables(características) seleccionadas por el usuario y retorna la predicción de la mejor ciudad.

---

<sup>12</sup> Registro: es un objeto que representa una colección de atributos.

### **Módulo de ranking de sitio turístico.**

En éste módulo se muestra el ranking de los sitios con mejores calificaciones. Inicialmente se llama a un webservice llamado rankingAtractivos(), el cual ejecuta función mejoresAtractivos(), esta función llama al objeto dataMiningRanking(), que ejecuta las funciones lineales discriminantes que utilizan las calificaciones del atractivo (atractivo, seguridad y limpieza) para dar como resultado el ranking de los sitios con mejores calificaciones según los criterios mencionados.

### **Módulo de recomendar nuevo sitio turístico**

Este módulo esta dividido en dos partes: recomendar nueva ciudad y recomendar nuevo atractivo.

Recomendar nueva ciudad: Para recomendar nueva ciudad se toman como datos de entrada los datos recolectados del formulario “Recomendar nueva ciudad”, éstos datos son validados para luego ingresarlos a la base de datos usando la función guardarCiudadRecomendada(), del objeto ciudad, asignándole el estatus de inactivo. Estos datos son automáticamente enviados vía e-mail al web master para que sean revisados.

Recomendar nuevo atractivo: una vez recolectada la información en el formulario de “Recomendar nuevo atractivo”, se llama a la función guardarNuevoAtractivo() del objeto atractivo, que almacena en la base la información del atractivo con un estatus de inactivo. Seguidamente se le envía un e-mail al Web master con los datos ingresados para que sean revisados.

### **Módulo de calificación de sitios.**

En el módulo de calificación de sitios se utiliza inicialmente la función guardarCalificación, que recibe como parámetros de entrada los datos ingresados en el formulario “Calificar Sitio Turístico”, estos datos pasan a ser cargados en el objeto revisión, el cual llama a su función insertarRevision(), que se encarga de almacenar los datos en la base.

## **5.3. RESULTADOS DE LA APLICACIÓN DE MÉTODOS DE MINERÍA DE DATOS**

Una vez realizada la implementación de los algoritmos antes descritos procedemos a indicar sus resultados:

### 5.3.1. RESULTADOS DE TÉCNICAS BAYESIANAS PARA EL MODELO PREDICTIVO

#### Primera fase - predicción del tipo de atractivo

Al evaluar las matrices de probabilidades generadas, usando el algoritmo de Naive Bayes se utilizan los datos ingresados por el turista para la predicción (actividades que desea realizar), con las cuales se generó el recorrido.

El resultado al final del recorrido es un vector de probabilidades de cada **tipo de atractivo** (la clase para este modelo), del cual se selecciona el que tenga la mayor probabilidad. Este dato sirve como una de las variables para el algoritmo de predicción de mejor ciudad.

El entrenamiento descrito, muestra los patrones de las diferentes actividades que generalmente se realizan en un determinado tipo de atractivo.

## **Segunda fase - predicción de mejor ciudad**

Al evaluar este algoritmo con los datos que se desea predecir, se genera un recorrido en los datos entrenados al aplicar el método, del cual se obtiene un vector con probabilidades de cada ciudad.

De este vector se escoge la ciudad que tenga mayor probabilidad, la cual sería la mejor opción para el turista, y el resto de resultados serían otras opciones que pueden ser tomadas en cuenta por el turista.

Las características de los atractivos y el tipo de atractivo, previamente predicho en la primera fase, se relacionan con las ciudades, estos son los patrones que resultan al final del modelo.

Con los patrones obtenidos logramos indicarle al usuario/turista/visitante el lugar más apropiado al que puede acudir para desde ahí desplazarse a los sitios turísticos donde pueda realizar las actividades de su preferencia.

Esta predicción se la muestra al usuario con la información más relevante de la ciudad y sus atractivos. Además de otras predicciones (ciudades).

Un ejemplo detallado de los resultados generados por la primera y segunda fase de la aplicación del algoritmo de Naive Bayes se encuentra disponible en el anexo 11.

### **5.3.2. RESULTADOS DE LA FUNCION LINEAL DISCRIMINANTE PARA EL MODELO DESCRIPTIVO.**

Los datos generados por las calificaciones de los usuarios hacia los diferentes atractivos, van a servir para realizar el ranking.

Al evaluar el modelo se obtienen los atractivos con mejores calificaciones de cada uno de los tres grupos existentes (calificación, concurrencia y limpieza).

Un ejemplo detallado de los resultados obtenidos al evaluar la función lineal discriminante se muestra en el anexo 10D.

#### 5.4. PRUEBAS REALIZADAS

Como se mencionó en el análisis del modelo en el capítulo 3, nuestras variables se enfocaron en las características y actividades del atractivo de acuerdo a las encuestas del anexo 2 y no en las variables con respecto al perfil del turista.

Mediante la aplicación del modelo predictivo se pretendía reconocer los patrones de las actividades que se realizan en un tipo de atractivo específico, dándonos como resultado la mejor ciudad para el turista.

Estos resultados fueron validados por 10 expertos en el área de turismo, entre ellos podemos destacar: Lcdo. Turismo Jorge Alberto Campoverde Valdez, Ex Director de Eventos del Hotel Colón Barcelo Miramar - Salinas; Master Carlos Ricaurte, Ex coordinador de Turismo - ESPOL; Master Paola Gálvez, Subsecretaria del Ministerio de Turismo del Ecuador; Lcdo. Gabriel Castro, área de eventos del Museo Municipal; Lcda. Daniela Isabel Rueda Valverde - TAME, Lcda. Mayra Echeverría; entre otros.

En la siguiente tabla se muestra se muestra como fueron divididos los usuarios que probaron el sistema:

Usuarios de pruebas del sistema			
NIVEL EXPERIENCIA	DESCRIPCION	RANGO DE EDAD	CANTIDAD
Expertos	Profesionales del área de turismo	26-34 años	10
Profesionales	Profesionales de diferentes áreas	30-45 años	15
Estudiantes universitarios	Estudiantes del área de turismo	20-26 años	7
Otros	Personas que viajan constantemente y conocen el país	24-82 años	20

**Tabla 5.7.** Usuarios de pruebas del sistema.

*Expertos:* se enfocaron en revisar los resultados de las predicciones de los sitios turísticos, evaluaron la predicción de atractivos por preferencias, obteniendo un alto nivel de aceptación, ya que en el 91% de los casos consiguieron como resultado el lugar que tenían en mente, a la hora de elegir las actividades y características del sitio.

*Profesionales:* los profesionales se orientaron a revisar la interfaz del sistema y el manejo de la información contenida, dando sus comentarios, sugerencias y mostrando un nivel de aceptación del sistema del 82%.

*Estudiantes Universitarios:* igual que los profesionales, se enfocaron en revisar la búsqueda de sitios por preferencia turística y la información presentada, obteniendo una confiabilidad del 96% en los resultados de



la predicción. Adicionalmente contribuyeron con el contenido de la información presentada, dando datos importantes de los sitios turísticos.

*Otros:* como personas conocedoras del país revisaron los resultados de la predicción y la información que muestra el sistema, determinando que posee información válida de los diferentes sitios turísticos y atractivos de la costa ecuatoriana, con una confiabilidad del 97.5%

Adicionalmente, se realizaron varias pruebas a partir de los patrones generados para verificar que todos los requerimientos hayan sido tomados en cuenta. A criterio de los expertos conocedores del tema del turismo los resultados generados se acercaban a sus expectativas con un alto grado de confiabilidad.

Finalmente se realizaron pruebas a cada módulo del sistema por separado, para corregirlos en el momento:

- **Módulo de búsqueda de sitios turísticos:** se verificó que los datos escogidos o ingresados generen correctamente el resultado deseado.

En la búsqueda por preferencias del turista, que es la parte principal del sistema que permite realizar la predicción de la ciudad en base a las actividades y características seleccionadas por el usuario, se comprobó que el algoritmo utilizado haya generado la mejor fidelidad de predicción.

- **Módulo de ranking de sitios turísticos:** se comprobó que los resultados sean mostrados correctamente al usuario, en base a los parámetros de: calificación, limpieza y concurrencia.
- **Módulo de recomendar nuevos sitios turísticos:** se validó que todos los campos ingresados de: recomendar nueva ciudad y recomendar nuevo atractivo; sean los adecuados, añadiéndose éstos a la base de datos para, luego de ser verificados por el administrador del sistema, se consideren en el proceso de minería.
- **Módulo de calificación de sitios:** se comprobó que se ingresen y escojan los datos necesarios de manera adecuada, validando cada uno de los campos. Este módulo sirve para posteriormente generar el ranking de los sitios turísticos.

Luego de haber realizado las pruebas descritas, notamos que los errores eran mínimos, y que básicamente se trataba de alguna variable mal declarada en la programación o por algún error en la consulta a la base de datos.

Con las pruebas generales a los usuarios del sistema, se pudieron comprobar errores de presentación, que fueron corregidos antes de que el sistema entre en funcionamiento, llegando a ser del completo agrado del usuario final.

## **CAPÍTULO 6**

# **PROPUESTA COMERCIAL DEL SISTEMA**

## 6.1. ANALISIS FODA

En el análisis FODA se da a conocer cuales son los factores que están a nuestro alrededor, que ayudan y a la vez afectan al sistema. Por cada punto se ha colocado varias características que servirán de base para una mejor organización en el sistema.

### **Fortalezas:**

- Actualmente no existe un sistema que brinde sugerencias al turista del lugar que debe visitar en base a sus preferencias.
- El turista puede consultar el ranking de los sitios más visitados de nuestra costa ecuatoriana.
- El sistema ofrece la información básica para que el turista pueda conocer el lugar adecuado según la búsqueda de sitios turísticos acorde a su perfil turístico.
- Se ofrece información variada de atractivos culturales, históricos, modernos, recreativos y otros.
- Sitio intuitivo para el usuario, con opciones funcionales y prácticas.

**Oportunidades:**

- No hay en el mercado otro sistema que de la oportunidad al turista de buscar el lugar ideal a visitar acorde a sus preferencias, basándose en las actividades que éste desee realizar y que además, brinde información puntual y útil de los lugares que pueda visitar en la costa ecuatoriana.
- Permitir la elección de distintos destinos turísticos por parte del turismo interno y externo.
- Conciencia y auge acerca de la importancia del sector turismo
- Nuevas tendencias turísticas mundiales que favorecen al turismo rural, el ecoturismo, el turismo de aventura y el turismo náutico.

**Debilidades:**

- Insuficiente cultura, falta de estrategias empresariales modernas y sensibilidad turística.
- Escasa promoción turística de varios sitios de la costa ecuatoriana (regional, nacional e internacional).
- Casi no hay conocimiento de las actividades de promoción al turismo de parte de las autoridades de la región.

- El sistema no abarca información turística de la sierra, oriente e islas galápagos.

**Amenazas:**

- Escaso apoyo de los gobiernos estatales y municipales para el desarrollo del turismo (infraestructura básica, promoción y capacitación turística).

## **6.2. PROPUESTA DEL VALOR COMERCIAL DEL SISTEMA**

Una vez realizado el análisis anterior para el sistema de búsqueda en línea de sitios turísticos, se pasa a indicar la justificación económica y el valor comercial del proyecto.

Este sistema está propuesto como un sistema de libre acceso sin costo alguno, para fomentar el desarrollo del turismo del Ecuador, el cual puede ser administrado por un ente gubernamental.

El nicho de mercado hacia el cual está orientado el sistema son las agencias de viajes, las mismas que pueden proporcionar al turista la

oportunidad de obtener mayor satisfacción al momento de decidir dónde viajar.

### 6.3. EVALUACIÓN DE COSTOS

El sistema está pensado para ser gratuito, por lo tanto no se incluyen los gastos en el desarrollo del mismo sino únicamente el costo de poner en marcha este proyecto, esto es: la infraestructura usada para el almacenamiento de las páginas Web y las licencias del software requerido.

La tabla 6.1 muestra el software utilizado en el desarrollo del proyecto y la tabla 6.2 muestra el software y hardware recomendado para implementar y poner en producción el sistema.

<b>ESTIMACIÓN DE COSTOS PARA DESARROLLO</b>	
Sistema Operativo Microsoft Windows XP Profesional + Internet Information Server, Inglés	\$ 130,00
Base de Datos Microsoft SQL 2000, Inglés	\$ 1.200,00
Microsoft Visual Studio.NET Versión Académica	\$ 0,00
<b>TOTAL</b>	<b>\$ 1.330,00</b>

**Tabla 6.1.** Software utilizado en el desarrollo del sistema. [11]



<b>ESTIMACIÓN DE COSTOS PARA PRODUCCIÓN DEL SISTEMA</b>	
Servidor Web (Sistema operativo incluido)	\$ 4.099,00
Servidor de Base de Datos (Sistema operativo y manejador de base de datos incluido)	\$ 6.116,00
Firewall	\$ 825,00
Herramientas varias de desarrollo	\$ 1.049,00
<b>TOTAL</b>	<b>\$12.089,00</b>

**Tabla 6.2.** Software y hardware recomendado para la implementación del sistema. [11]

Dado que los costos de poner en producción el sistema fueron de \$12.089,00 y el software utilizado en el desarrollo de \$1.330,00; cabe acotar que las herramientas de software utilizadas en el desarrollo del sistema ya estaban a disposición, reduciendo los costos del desarrollo a \$ 0, 00, haciendo viable el proyecto.

A continuación se detallan las características del equipo necesario para la instalación del servidor Web y del servidor de base de datos recomendados en la tabla 6.2:

*Servidor Web:*

- 2 Procesadores Intel Xeon de 3.2GHz, 2MB caché.
- 1 Gigabyte de Memoria DDR2, 400MHz.
- Windows Server 2003.

- Tarjeta RAID SCSI<sup>13</sup>.
- 3 discos duros de 73 Gigabytes RPM SCSI.

*Servidor de base de datos:*

- 2 Procesadores Intel Xeon de 2.8GHz, 2MB caché.
- 1 Gigabyte de Memoria DDR2, 400MHz.
- Windows Server 2003.
- Tarjeta RAID SCSI.
- 3 discos duros 73 Gigabytes RPM SCSI.
- Licencia Microsoft SQL Server 2000 Standard.

#### **6.4. REQUERIMIENTOS MÍNIMOS DEL SISTEMA**

El sistema de búsqueda en línea de sitios turísticos puede funcionar sin problema alguno bajo la siguiente plataforma de hardware y software:

Hardware recomendado para acceder al sistema:

- Procesador Intel Pentium III.
- 256 MB de memoria RAM.
- Tarjeta de red 10/100 Mbps.

---

<sup>13</sup> SCSI (Small Computer System Interface) es una interfaz estándar para la transferencia de datos entre periféricos en el bus del computador

Software recomendado para acceder al sistema:

- Sistema operativo Microsoft Windows XP.
- Navegador web Internet Explorer 6.0 o superior
- Navegador web Mozilla Firefox 1.5

## CONCLUSIONES

- En base de las comparaciones realizadas en el capítulo 3, se puede decir que el algoritmo de Naive Bayes es eficiente, ya que utilizado en el sistema generó un mejor grado de confiabilidad en la predicción, considerando todas las variables del sistema.
- Aunque se dice que la minería de datos viene asociada a grandes cantidades de datos, por la carencia de dicha base, se creó un repositorio moderado de datos, previamente depurados, con el cual el sistema intenta demostrar la calidad de los datos predichos, a través de una muestra, generando conocimiento predictivo en un tiempo razonable para la toma de decisiones.
- El proceso de minería de datos no se aplica necesariamente a grandes cantidades de datos, ya que inicialmente se realiza un proceso de selección, limpieza y transformación de dichos datos, obteniendo únicamente aquellos que permitan generar el modelo adecuado, siendo este un conjunto reducido del total existente en la base de datos.

- De acuerdo a las encuestas realizadas no fue necesario incluir variables con respecto del usuario sino variables con respecto del atractivo para generar la predicción como se demostró en el anexo 2.
- Los resultados obtenidos en el sistema poseen un alto grado de confiabilidad, acorde a las pruebas realizadas con los usuarios expertos en el área de turismo.
- Los objetivos delineados al inicio de esta investigación, en lo que concierne al desarrollo de este proyecto fueron cumplidos, incluso no fue necesaria la adquisición de software o hardware adicional ya que se emplearon recursos tecnológicos que ya estaban a nuestra disposición.

## RECOMENDACIONES

- Aunque el sistema está pensado solamente para la costa ecuatoriana se puede ampliar su uso para las regiones: sierra, oriente e insular del país.
- Crear un plan de mantenimiento para tener siempre al día el inventario turístico (nuestra base de conocimientos) de sitios y atractivos.
- Buscar el respaldo de instituciones privadas y públicas para promocionar el sistema.
- Sugerimos utilizar el algoritmo de Naive Bayes para resolver tareas predictivas similares donde existen múltiples variables, por su simplicidad en la implementación y el tiempo necesario que se toma para inducir el modelo.
- Este sistema podría constituirse en una herramienta que puede servir de base al turismo y convertirse, siendo aplicado adecuadamente, en un canal de promoción de la actividad turística de la costa ecuatoriana.

## REFERENCIAS BIBLIOGRÁFICAS

- [1] Ministerio de Turismo, “Ecuador y las Islas Galápagos”, <http://www.vivecuador.com/html2/esp/home.htm>
- [2] Hotels.com, L.P, “El Especialista en Hoteles”, <http://sa.hoteles.com/hotel-htdocs/hoteles.com/sa/95086index.htm>
- [3] Renta de autos en-línea, [www.avis.com](http://www.avis.com), [www.hertz.com](http://www.hertz.com), [www.budget.com](http://www.budget.com)
- [4] Compra de boletos aéreos en-línea, [www.tame.com.ec](http://www.tame.com.ec), [www.continental.com](http://www.continental.com)
- [5] Raúl Jimmy Álvarez Guale, “Construcción de una base de datos sobre turismo en la sierra y oriente ecuatoriano y su difusión en Internet”, Tesis de Grado, ESPOL, 1999, Capítulo 1, Pág. 28, <http://www.cib.espol.edu.ec/bivir/tesis.asp?tco=2B425E565A575E4F4F4F4F4F4F4F4F4F>
- [6] José Hernández Orallo, Ma. José Ramírez Quintana, César Ferri Ramírez “Introducción a la minería de datos”, Primera Edición, Pearson Prentice Hall, México, 2004.
- [7] Álvaro Clemente, Pablo López, José Penado, “ Minería de Datos y Minería Web”, Departamento de Ingeniería en Sistemas Telemáticos, Universidad Politécnica de Madrid, 2005, <http://www.gsi.dit.upm.es/~gfer/ssii/trabajos2005/MineriadeDatosAlvaroClemente.pdf>
- [8] Michael J. A. Berry, Gordon S. Linoff, “Data Mining Techniques”, Segunda Edición, Wiley, 2004.
- [9] “Importancia de la Toma de Decisiones en las distintas áreas según Stoner y Freeman”, <http://www.monografias.com/trabajos17/modelo-de-decisiones/modelo-de-decisiones.shtml>
- [10] Gastón Quirque, “Introducción al SQL Server 2000”, Introducción a SQL Server 2000.ppt.
- [11] Sitio Web de la compañía Dell, [www.dell.com](http://www.dell.com)

- [12] Enrique Bonsón Ponte, Tomás Escobar Rodríguez, M<sup>a</sup> Del Pilar Martín Zamora “*Sistemas de Inducción de Árboles de Decisión: Utilidad en el Análisis de Crisis Bancarias*”, Grupo de Inteligencia Artificial en Contabilidad y Administración de Empresas, Universidad de Huelva <http://ciberconta.unizar.es/Biblioteca/0007/arboles.html>



# **ANEXOS**

## **ANEXO 1**

### **TURISMO EN EL ECUADOR**

Históricamente, hacer turismo constituyó un privilegio de las clases adineradas, es decir, de la clase social constituida por los grupos o estratos de gran poder económico. Muchas de estas personas practicaban los viajes, además del placer que de ello se derivaba, como un acto expreso de sobresalir notoriamente como persona importante. Hoy en día todo ha cambiado a través de la historia y el turismo se ha convertido en una necesidad y en un placer accesible a mayor número de personas y estratos sociales.

El turismo es una actividad cuyo protagonista es el hombre, por lo que, al desarrollarse en el campo personal, cada quien tiene su interpretación, la cual depende de: sus vivencias, sus deseos, sus gustos, sus motivos, su cultura y su idioma; por lo tanto su definición desde el punto de vista general se torna dificultosa y a veces contradictoria

La palabra turismo, etimológicamente<sup>14</sup> considerada se deriva de la palabra latina " **tornus** " que quiere decir vuelta o movimiento y la cual adoptaron los ingleses, franceses y españoles con ese sentido, aplicándola al "**turismo**". Por eso, hoy cuando se ofrece un viaje turístico, o sea un movimiento de personas que se trasladan o viajan por placer, se les invita a participar de un " tour " a determinado destino.

Podemos definir el turismo de una forma más simple y común diciendo que "es la tendencia natural del ser humano a cambiar de sitio para beneficiarse de las bondades de otros lugares distintos de aquel donde usualmente vive y descansar en un ambiente agradable, para esparcimiento y recreación"<sup>15</sup>, y porque no también, el movimiento de humanos para intercambiar: conocimientos, cultura, aventuras, credos, idiomas.

La costa ecuatoriana se encuentra bañada por el océano Pacífico, está conformada por cinco provincias, de norte a sur, que son: Esmeraldas, Manabí, Guayas, Los Ríos y El Oro.

---

<sup>14</sup> La palabra etimología viene del latín "*etimología*" y este del griego "*etymologia*". Está formado de *etymos* (verdadero) y *logos* (palabra). Es decir, significa "verdadero origen de las palabras". La "*Étimo*" deriva de la palabra etimología y se refiere a la raíz de un vocablo.

<sup>15</sup> Rafael Ramírez Castellano es Ingeniero Civil (Universidad de los Andes). Actualmente es Presidente de C.A. Constructora RAMCAS y de INVERSIONES PARA TURISMO C.A (Hotel La Pedregosa)

La costa está formada por tres ecosistemas principales: los bosques lluviosos tropicales del norte, las sabanas tropicales del centro y del sudoeste, y el bosque seco de la franja peninsular occidental y meridional. A lo largo del litoral costero se distinguen dos ecosistemas adicionales caracterizados por sus comunidades animales y vegetales: las entrantes de manglar, las playas y acantilados conocidos por su peculiar formación rocosa.

La temperatura promedio que prevalece en la costa es de 22°C. El invierno dura de diciembre a mayo y el verano de junio a diciembre. La corriente cálida conocida como corriente del Niño provoca un aumento en los índices de pluviosidad entre los meses de enero y mayo.

Su territorio está formado por: llanuras fértiles, colinas, cuencas sedimentarias y elevaciones de poca altitud. Por su territorio corren ríos que parten desde los Andes hasta llegar al océano Pacífico. Sus cinco provincias cuentan con playas y balnearios muy atractivos para el turista. Se destacan las de: Esmeraldas, Manabí y Guayas. En esta zona se encuentra la red fluvial más extensa del país. Se trata de la cuenca del río Guayas, que tiene cerca de doce afluentes junto a las poblaciones de: Daule, Babahoyo, Macul, Puca, Paján y Colimes. Ecuador tiene 640 kilómetros de costa.

De norte a sur presenta hermosas playas que ofrecen descanso y diversión, y cuentan con una importante infraestructura hotelera. La región se caracteriza por un clima cálido y seco al sur, y tropical húmedo al norte. Varios Parques Nacionales como Machalilla y Manglares-Churute, ofrecen la posibilidad de realizar ecoturismo. En Guayaquil, el puerto principal, se encuentra el aeropuerto internacional José Joaquín de Olmedo.

Ésta es una de las tres regiones naturales en las que se divide el Ecuador continental, se extiende entre la cordillera de los Andes y el océano Pacífico. Con escasa diferencia de altitud, la región costanera está dotada de uniformidad de climas y de paisajes, sin embargo, se encuentran variaciones desde el bosque tropical en el norte hasta las desérticas llanuras en el sur.

La llamada “Costa Azul” ecuatoriana ofrece hermosos balnearios con extraordinarias playas, arena limpia y brillante sol, que permiten al turista gozar de un clima agradable, de deportes acuáticos, de incomparable pesca y de toda clase de entretenimientos en confortables hoteles.

Los balnearios del Ecuador brindan todas las comodidades de vivienda y turismo. Salinas, Ayangue, Puerto López, Súa, Manta, Bahía de Caráquez, Esmeraldas, Playas, entre otros, son de gran seguridad.

Provincia de Esmeraldas, cuenta con bellas y hermosas playas de arena suave y mar tibio como: Atacames, Súa, Tonsupa, Playa Ancha, Tonchigüe, Camarones y Same.

Esmeraldas es conocida como "la ciudad del Sol" por estar alumbrada por un hermoso sol tropical. Aquí se encuentran el museo del Banco Central del Ecuador con muestras de la cultura Tolita y muchos otros atractivos típicos como el folclore negro y su exquisita gastronomía.

Hacia el sur está la encantadora Provincia de Manabí, sus playas invitan al turista a: disfrutar, descansar y gozar del sol, la arena y el mar. Entre las playas más conocidas están: Cojimíes, Pedernales, Camarones, Canoa, San Vicente, Bahía de Caráquez, San Clemente, San Jacinto, Crucita, Jaramijó, San Mateo, San José, Puerto Cayo, Machalilla, Los Frailes, Puerto López y Salando.

El puerto de Manta, es una de las mejores áreas para la pesca deportiva en alta mar, es importante resaltar el parque nacional Machalilla, que

comprende un territorio continental e insular con un bosque seco tropical y zonas bajas costeras con playas y acantilados; Portoviejo "La ciudad de los Reales Tamarindos" es una ciudad de gran movimiento comercial, industrial y artesanal.

Provincia del Guayas, ofrece al turismo sus maravillosas playas y balnearios frecuentada por los amantes de la pesca de profundidad y los deportes acuáticos. A pocos kilómetros de Guayaquil, en la península de Santa Elena se encuentra el moderno balneario de Salinas, centro de atracción turística nacional e internacional, que ofrece servicios turísticos de primera clase: hoteles, casinos, restaurantes, y altos condominios de hermosa arquitectura, además de su club de yates con modernas instalaciones. Desde Salinas se puede visitar otras playas cercanas como: Ballenita, Punta Carnero, y La Puntilla. El balneario de Playas, muy cerca se encuentra el pequeño puerto pesquero de Posorja. Otras playas de la provincia son: Manglar Alto, Montañita, Palmar, y Punta Blanca.

La ciudad de Guayaquil es el pulmón: económico, comercial, y financiero del país. Situada en el golfo de Guayaquil junto al importante río Guayas; parque El Malecón desde el cual pueden observarse las gabarras que navegan por el majestuoso río; el monumento

conmemorativo a la célebre entrevista de los libertadores Simón Bolívar y San Martín en 1.822; la Torre del Reloj, de estilo morisco y los importantes edificios coloniales del Municipio y el de la Gobernación.

El tradicional barrio Las Peñas, situado al pie del cerro del Carmen, conserva la muestra de la historia colonial de la ciudad. En el cerro del Carmen, el Sagrado Corazón bendice la ciudad. El cementerio General con: sus esculturas, sus jardines y sus mausoleos de estilo greco romano. Guayaquil cuenta con interesantes museos: Banco Central del Ecuador, Casa de la Cultura, Banco del Pacífico, Museo Municipal, entre otros.

La Provincia de Los Ríos, denominada así por la cantidad de ríos que la cruzan. Posee grandes extensiones de campos cultivados de banano, café, cacao y arroz. Entre los principales centros turísticos están: las ciudades de Babahoyo y Quevedo de gran actividad agropecuaria y comercial; el cerro Cacharí y su curiosa estructura, el estero de Dimas, la parroquia Barreiro, con todo su paisaje esplendoroso.

La Provincia de El Oro, poseedora de la mejor producción en el mundo de banano y camarón. Cuenta con atractivos turísticos de importancia y aceptación para quién la visita, tales como: las minas de oro en Zaruma;



las más bellas especies de orquídeas en Piñas, numerosos balnearios y playas; y el único bosque Petrificado de Puyango que limita con Loja. Se destacan, además de las tradicionales, la Feria Mundial del Banano y el festival de Café. Su comida típica a base de plátano verde y camarones. La costa en general cuenta con gente amable y cordial, playas amplias en un ambiente tropical, en donde saboreará una gastronomía típica a base de productos del mar

## ANEXO 2

### ENCUESTA REALIZADA PARA RECONOCER ATRIBUTOS DE ATRATIVOS TURÍSTICOS

La siguiente encuesta es para conocer los atributos de los lugares turísticos que usted desee visitar. Agradecemos su gentil colaboración.

Edad		
Sexo	M	F

Para las siguientes preguntas, por favor imagine un lugar que usted desee visitar.

1. Describa los atributos que tiene el lugar que usted imaginó
2. ¿Cuáles son las actividades que usted desea practicar cuando realiza turismo?
3. ¿Es necesario que tenga seguridad el lugar que usted desea conocer?

Si \_\_\_

No \_\_\_

4. ¿Qué tan importante es tener un lugar donde comer cuando usted está visitando un sitio turístico? (califique de 1 a 3)

Mucho	Poco	Indiferente
1	2	3

5. ¿Usted requiere que el lugar que visita tenga parqueadero para su comodidad?

Si \_\_\_

No \_\_\_

### RESULTADOS DE LA ENCUESTA

La encuesta fue realizada a 75 personas: 34 hombres y 41 mujeres; al realizar la encuesta pudimos notar que las personas se enfocaban principalmente en ciertas características del lugar que desean visitar y en las actividades que pueden realizar en dicho lugar. A continuación se muestran los resultados obtenidos según cada una de las preguntas realizadas:

SEXO	
Masculino	45%
Femenino	55%

EDAD	
18-25	27%
26-32	45%
33-50	28%

**1. ¿Describe los atributos que tiene el lugar que usted imaginó?**

Atributos del lugar			
Atributos	Hombres	Mujeres	Total
Limpieza	94,12%	85,37%	89,74%
Seguridad	88,24%	95,12%	91,68%
Poco concurrido	50,00%	85,37%	67,68%
Comodidad	41,18%	48,78%	44,98%
Lugar para comer	88,24%	100,00%	94,12%
Tranquilo	50,00%	80,49%	65,24%
Áreas Verdes	35,29%	48,78%	42,04%
Que no tenga plantas	8,82%	7,32%	8,07%
Que sea Playa	58,82%	36,59%	47,70%
Clima templado	5,88%	9,76%	7,82%
Otros	17,65%	7,32%	12,48%

Dentro de los resultados arrojados podemos observar que los atributos que obtuvieron mayor porcentaje han sido: limpieza, seguridad, concurrencia, lugar para comer; demostrando que son muy relevantes para los turistas, por lo que han sido considerados como variables del sistema.

**2. ¿Cuáles son las actividades que usted desea practicar cuando realiza turismo?**

<b>Actividades a realizar</b>			
	<b>Hombres</b>	<b>Mujeres</b>	<b>Total</b>
Admirar fauna	5,88%	12,20%	9,04%
Admirar flora	8,82%	21,95%	15,39%
Alquilar cabañas	2,94%	7,32%	5,13%
Bañarse en el mar	11,76%	19,51%	15,64%
Cantar karaoke	8,82%	19,51%	14,17%
Comprar artesanías	2,94%	24,39%	13,67%
Degustar platos típicos	14,71%	26,83%	20,77%
Descansar en hamacas	5,88%	7,32%	6,60%
Escuchar música en vivo	17,65%	2,44%	10,04%
Ingerir bebidas alcohólicas	23,53%	9,76%	16,64%
Ir a bares	8,82%	12,20%	10,51%
Ir a discotecas	17,65%	17,07%	17,36%
Ir de compras	0,00%	31,71%	15,85%
Jugar basket	8,82%	7,32%	8,07%
Jugar fútbol	20,59%	0,00%	10,29%
Jugar volley	23,53%	4,88%	14,20%
Montar a caballo	20,59%	9,76%	15,17%
Motos acuáticas	20,59%	7,32%	13,95%
Nadar	29,41%	21,95%	25,68%
Observar orquídeas	0,00%	19,51%	9,76%
Observar Peñascos	17,65%	0,00%	8,82%
Ver piqueros de patas azules	2,94%	24,39%	13,67%
Pasear a caballo en la playa	26,47%	2,44%	14,45%
Practicar buceo	29,41%	0,00%	14,71%
Practicar fútbol de playa	20,59%	7,32%	13,95%
Practicar motocross	8,82%	2,44%	5,63%
Practicar senderismo	17,65%	9,76%	13,70%
Recolectar crustáceos	0,00%	14,63%	7,32%
Tomar el sol	26,47%	31,71%	29,09%
Visitar cafeterías	29,41%	12,20%	20,80%
Visitar comunidades indígenas	5,88%	7,32%	6,60%
Visitar haciendas	23,53%	12,20%	17,86%
Visitar Iglesias	8,82%	19,51%	14,17%
Visitar monumentos históricos	14,71%	14,63%	14,67%
Visitar museos	17,65%	17,07%	17,36%
Visitar parques nacionales	2,94%	21,95%	12,45%
Visitar reservas de bosques	11,76%	14,63%	13,20%
Visitar ríos	17,65%	7,32%	12,48%
Visitar ruinas arqueológicas	14,71%	0,00%	7,35%
Otras	47,06%	31,71%	39,38%

Los resultados obtenidos con esta pregunta demuestran que las actividades que desean realizar los turistas son muy variadas y no dependen del sexo o edad de las personas, por lo que fueron consideradas todas las actividades posibles que se pueden realizar en los diferentes atractivos.

**6. ¿Es necesario que tenga seguridad el lugar que usted desea conocer?**

<b>Seguridad</b>			
	<b>Hombres</b>	<b>Mujeres</b>	<b>Total</b>
<b>Si</b>	58,82%	85,37%	72,09%
<b>No</b>	41,18%	14,63%	27,91%

Estos resultados demuestran que la seguridad es un factor importante a la hora de seleccionar el lugar a visitar, por lo que se consideraron los siguientes niveles: alta, media, baja y no interesa; para brindar mayor flexibilidad.

**7. ¿Qué tan importante es tener un lugar donde comer cuando usted está visitando un sitio turístico? (califique de 1 a 3)**

<b>Lugar para Comer</b>			
	<b>Hombres</b>	<b>Mujeres</b>	<b>Total</b>
<b>1</b>	58,82%	73,17%	66,00%
<b>2</b>	29,41%	12,20%	20,80%
<b>3</b>	11,76%	14,63%	13,20%

La tabla anterior muestra la importancia de la existencia de un lugar para comer en un atractivo turístico; para ésta variable se consideraron los valores: si, no y no interesa.

8. **¿Usted requiere que el lugar que visita tenga parqueadero para su comodidad?**

<b>Parqueadero</b>			
	<b>Hombres</b>	<b>Mujeres</b>	<b>Total</b>
<b>Si</b>	64,71%	85,37%	75,04%
<b>No</b>	35,29%	14,63%	24,96%

Los resultados de la tabla anterior demuestran que, para una gran cantidad de visitantes, es muy importante que el lugar que visitan tenga parqueadero en donde puedan dejar su vehículo, esto nos permitió considerar el parqueadero como una variable del sistema con 3 niveles: si, no y no interesa.

## ANEXO 3

### CLASIFICACIÓN DE ACTIVIDADES TURÍSTICAS

CATEGORÍA	DESCRIPCIÓN
Deportes	Deportes tradicionales que se practican
Deportes Extremos	Deportes en los cuales se incluye alto riesgo físico
Diversión Nocturna	Discotecas, bares, peñas, etc. Se requiere que la persona sea mayor de edad para entrar a estos lugares
Cultura y Ciencia	Visita de Sitios Culturales, históricos y científicos
Ecoturismo	Turismo Ecológico
Gastronomía	Lugares Gastronómicos
Aventura	Turismo Aventura (camping)
Recreación	Lugares de Esparcimiento
Otras Actividades	Actividades varias que se pueden realizar en un atractivo



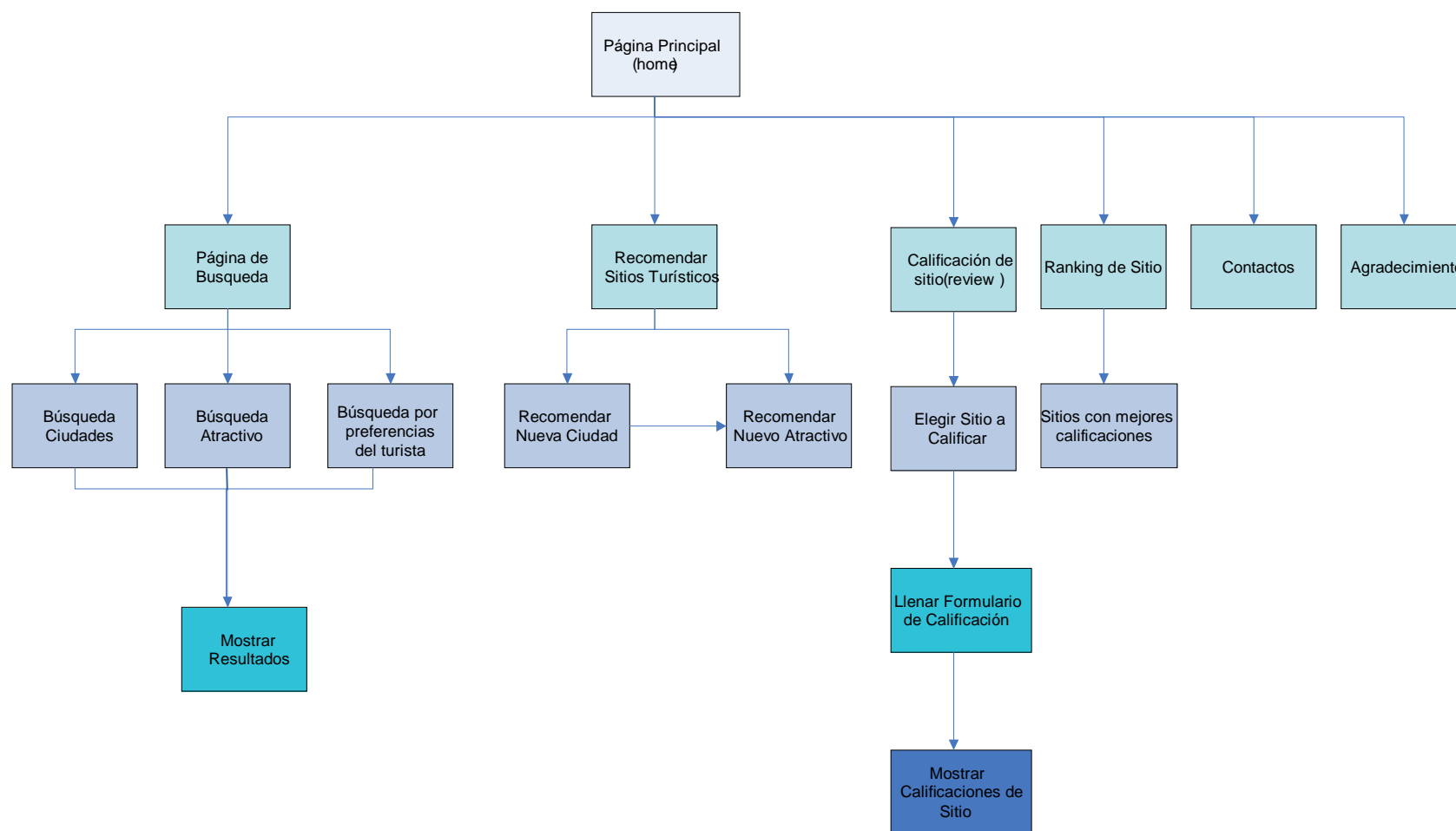
## ANEXO 4

### TIPOS DE ATRACTIVOS TURÍSTICOS

IDENTIFICADOR EN LA BASE DE DATOS	TIPO ATRACTIVO	DESCRIPCIÓN
1	Campo	Esta categoría comprende: haciendas, llanuras, montañas, cerros, cascadas, fincas y granjas
2	Balneario	Incluye playas, ríos, esteros, lagos, y lagunas
3	Cultural	Esta categoría comprende: museos, iglesias, monumentos, arte, teatro, ruinas, y parques arqueológicos.
4	Parques y Malecones	Esta categoría comprende los diferentes parques rurales y malecones
5	Comerciales	Esta categoría comprende los diferentes centros comerciales, mercados artesanales y tiendas donde se realiza comercio
6	Reservas Ecológicas	En este lugar podemos encontrar: manglares, bosques protectores, reservas naturales y zoológicos.

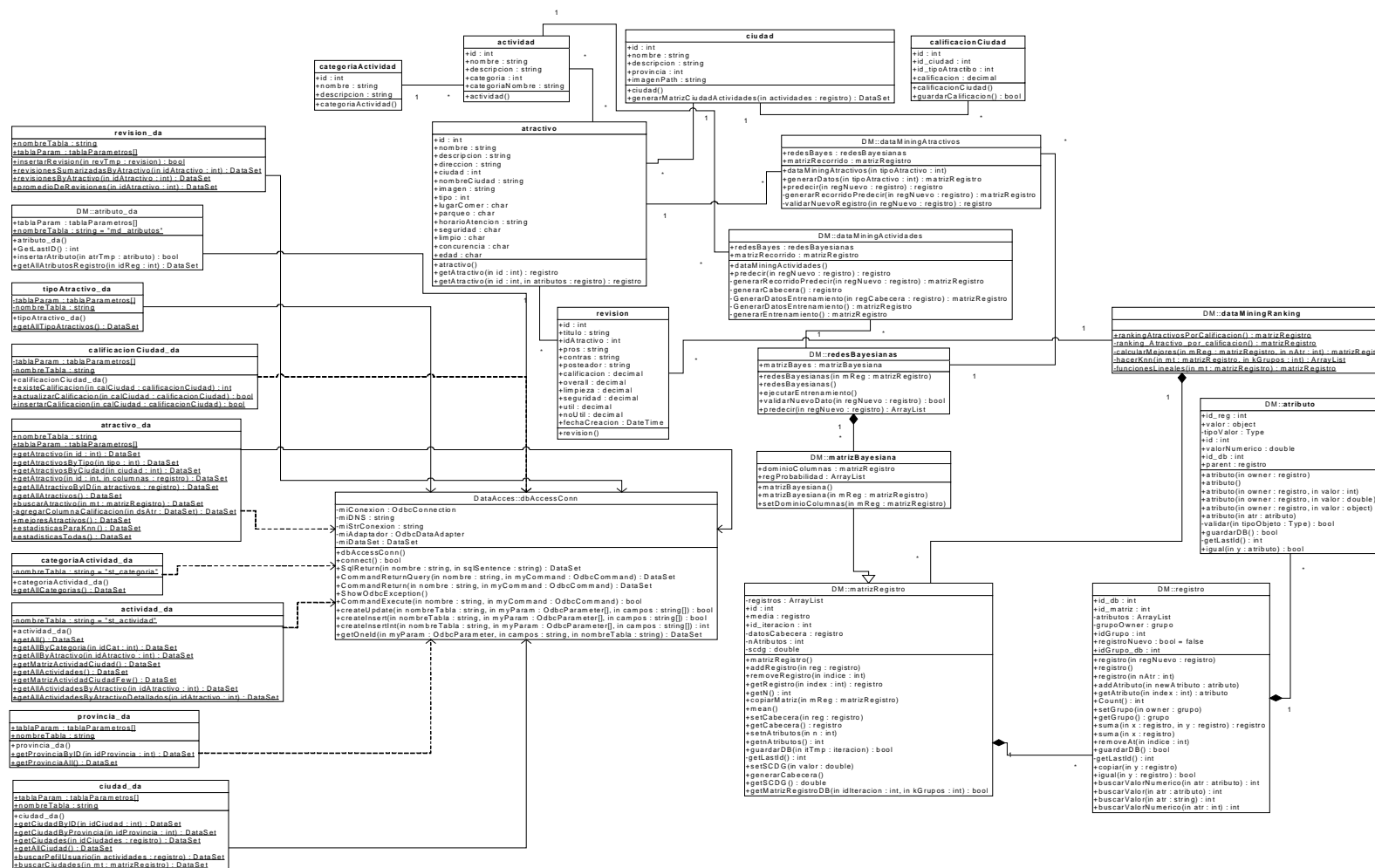
## ANEXO 5

### FLUJO DE VENTANAS



# ANEXO 6

## DIAGRAMA DE CLASES



## **ANEXO 7**

### **ESCENARIOS DE LOS CASOS DE USO DEL SISTEMA**

#### **Caso de Uso 1: Ingresar al sistema**

##### **Escenario 1.1**

##### **Ingreso exitoso del sistema**

##### **Supuestos**

- Dirección del sitio web disponible.
- Información de la base de datos está arriba y en línea.

##### **Salidas:**

- Usuario ingresa al sistema.
- Usuario navega dentro del sistema satisfactoriamente.

##### **Escenario 1.2**

##### **Ingreso no exitoso del sistema**

##### **Supuestos**

- Dirección del sitio web no disponible.
- Base de datos no posee información.

##### **Salidas:**

- Usuario no ingresa al sistema.
- Usuario no puede navegar dentro del sistema.

## **Caso de Uso 2: Búsqueda por perfil de usuario**

### **Escenario 2.1**

**Sistema realiza búsqueda de sitios turísticos exitosamente.**

#### **Supuestos:**

- Usuario/turista/visitante desea conocer sitios turísticos de la costa.
- Usuario/turista/visitante decide ingresar a la búsqueda por perfil.
- Usuario/turista/visitante escoge adecuadamente los campos necesarios de las características del lugar y las actividades a realizar para predecir las ciudades dentro de las cuales puede encontrar los atractivos turísticos.

#### **Salidas:**

- El sistema muestra la ciudad que mejor se perfila a sus necesidades, en base a las características que el usuario escoge satisfactoriamente.

### **Escenario 2.2**

**Sistema no puede realizar búsqueda de sitios turísticos exitosamente.**

#### **Supuestos:**

- Usuario/turista/visitante no decide ingresar a la búsqueda por perfil.

- Usuario/turista/visitante no escoge adecuadamente o completamente los campos necesarios para realizar la búsqueda.

**Salidas:**

- El sistema no muestra búsqueda alguna.

**Caso de Uso 3: Calificación y ranking de sitios turísticos****Escenario 3.1**

**Sistema actualiza calificación y ranking de sitios turísticos exitosamente.**

**Supuestos:**

- Usuario ha ingresado al formulario de calificación de sitios turísticos.
- Usuario llena los datos de calificación correctamente.
- Usuario ya ha visitado y conoce los sitios turísticos disponibles.

**Salidas:**

- El sistema muestra ranking de los mejores atractivos calificados (sitios turísticos) junto con las puntuaciones realizadas por el usuario satisfactoriamente.
- Contador de calificación para ranking de sitios es incrementado.

### **Escenario 3.2**

**Sistema no actualiza calificación y ranking de sitios turísticos exitosamente.**

**Supuestos:**

- Usuario no ha ingresado correctamente el formulario de calificación de los sitios.
- Sitio turístico no está disponible.

**Salidas:**

- Contador de calificación para ranking de sitios no es incrementado.
- El sistema no muestra actualización de ranking de los sitios calificados.

### **Caso de Uso 4: Recomendar sitios turísticos**

#### **Escenario 4.1**

**Usuario ingresa recomendación de sitios turísticos exitosamente.**

**Supuestos:**

- Usuario/turista/visitante decide ingresar a la opción de recomendar nuevo sitio turístico.
- Usuario/turista/visitante ingresa correctamente los datos del formulario para recomendar una nueva ciudad o un nuevo atractivo.

**Salidas:**

- El sistema almacena exitosamente todas las recomendaciones ingresadas por el usuario.

**Escenario 4.2**

**Usuario no puede ingresar recomendación de sitios turísticos exitosamente.**

**Supuestos:**

- Usuario/turista/visitante decide ingresar a la opción de recomendar nuevo sitio turístico.
- Usuario no ha ingresado todos los datos para una nueva recomendación

**Salidas**

- El sistema no ingresa la nueva recomendación.
- El sistema muestra un mensaje de recomendación no exitosa.



## **ANEXO 8**

### **MANUAL DE USUARIO**

El sistema de búsqueda en línea de sitios turísticos, es un sitio Web que permite a los usuarios determinar el lugar más apropiado que deben visitar, según las actividades que deseen realizar y las características del sitio.

Adicionalmente, los usuarios pueden informarse de cuáles son los atractivos más recomendados y obtener valiosa información de los encantos turísticos de la costa ecuatoriana.

#### **Requerimientos de Hardware**

- Procesador Intel Pentium III.
- 256 MB de memoria RAM.
- Tarjeta de red 10/100 Mbps.

#### **Requerimientos de Software**

- Sistema operativo Microsoft Windows 2000/XP.
- Navegador web (Internet Explorer 6.0/Mozilla Firefox 1.5 o superiores)

## Ingreso al Sistema

Para hacer uso del sistema, debe conectarse a Internet utilizando uno de los navegadores web soportados y acceder a la dirección:

[www.turismocosta.com.ec](http://www.turismocosta.com.ec)

Una vez que el usuario haya ingresado al sitio web del sistema, visualizará la pantalla de bienvenida. En esta ventana se dispone de un menú desplegable que le permite al usuario interactuar con el sistema y se muestran los lugares más recomendados a visitar.



El menú está compuesto por:

- Búsqueda de sitios turísticos,
- Recomendar nuevo sitio turístico,
- Calificación de sitios turísticos,
- Ranking de los sitios turísticos
- Contáctenos.

### **Buscar ciudades de la costa**

Si el usuario desea realizar una búsqueda de las ciudades de la costa ecuatoriana debe dirigirse a “Búsqueda de Sitios Turísticos” y escoger la opción “Búsqueda de Ciudades”, para realizar la búsqueda debe poner parte o todo el nombre de la ciudad de la que desea información y, si lo desea, seleccionar la provincia en la que desea buscar dicha ciudad.

Como se indica a continuación:



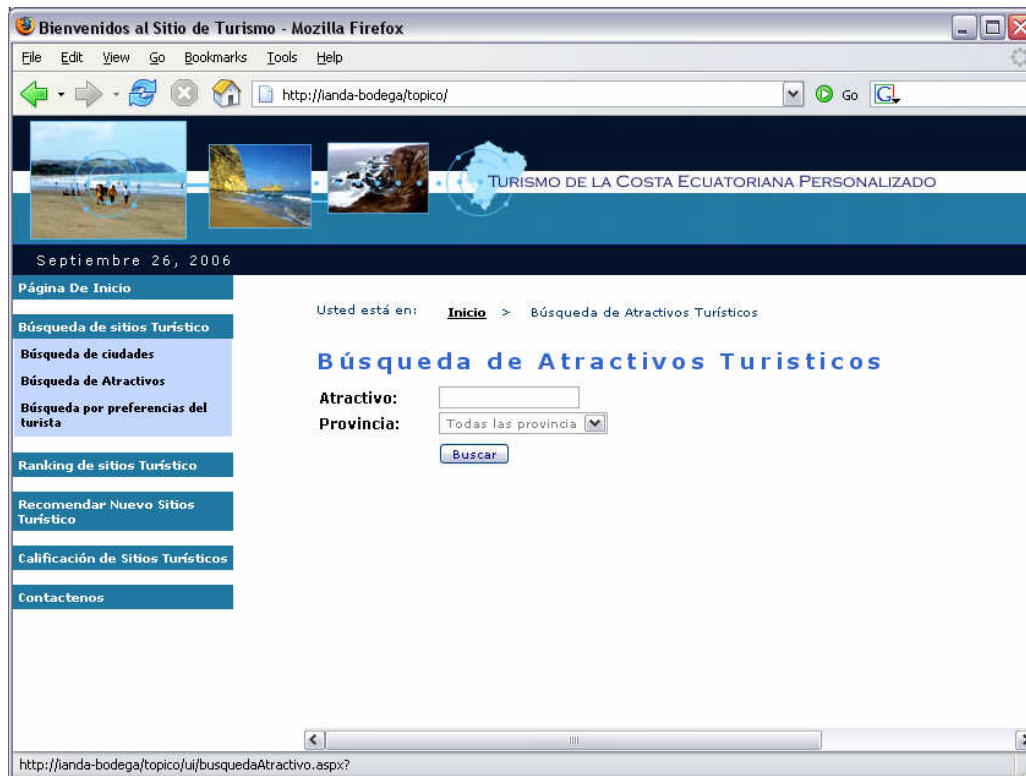
Con esto obtendrá como resultado todas aquellas ciudades que contengan, en su nombre, lo que el usuario haya escrito, tal como se muestra a continuación:



Para obtener mayor información de la ciudad y sus atractivos, puede dar click sobre el nombre o sobre la imagen asociada a ésta.

### Buscar atractivos turísticos

Si desea buscar un atractivo en especial debe dirigirse a: "Búsqueda de Sitios Turísticos" y escoger la opción "Búsqueda de Atractivos".



En esta ventana debe poner parte o todo el nombre del atractivo que desea buscar. Adicionalmente, tiene la opción de elegir la provincia y la ciudad a la que pertenece el atractivo del cual desea obtener mayor información.

Una vez que haya ingresado los datos necesarios se mostrará una lista de los atractivos que coincidan con la búsqueda realizada, indicando además una breve descripción del mismo.

Bienvenidos al Sitio de Turismo - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

http://ianda-bodega/topico/

TURISMO DE LA COSTA ECUATORIANA PERSONALIZADO

Septiembre 26, 2006

Página De Inicio

Búsqueda de sitios Turístico

Búsqueda de ciudades

Búsqueda de Atractivos

Búsqueda por preferencias del turista

Ranking de sitios Turístico

Recomendar Nuevo Sitios Turístico

Calificación de Sitios Turísticos

Contactenos

Usted está en: Inicio > Búsqueda de Atractivos Turísticos.

### Búsqueda de Atractivos Turísticos

Atractivo:

Provincia:

Ciudad:

Atractivo	Calificacion	Ciudad	Descripcion	Imagen
<b>Zoológico El Pantanal</b>	☆☆☆☆☆	<b>Guayaquil</b>	Este zoológico es un centro de rescate y conservación de nuestra fauna donde podrá disfrutar del color y vida de la naturaleza y del placer de observar cientos de animales, en su mayoría propios del Ecuador, en un ambiente lleno de verdor y paz.	

Done

Dando click sobre el nombre del atractivo o sobre su imagen obtendrá mayor información de las características y las actividades del atractivo turístico encontrado.

### El atractivo Seleccionado es:

Atractivo:	<b>Playa de Atacames</b>
Descripción:	Cuenta con una amplia infraestructura hotelera. A lo largo del malecón se ubican los principales restaurantes, con lo mejor de la comida típica de la región. Junto a la playa, numerosos bares ofrecen bebidas típicas a base de agua de coco, principalmente. También hay un paseo con artesanías, accesorios para el cuerpo y objetos trabajados en coral.
Dirección:	Esta ubicada a 30 Km al suroeste de la ciudad de Esmeraldas. se puede llegar con buses directos desde Quito, Santo Domingo y Esmeraldas.
Lugar Para Comer:	Si
Parqueo:	Si
Nivel de Seguridad:	Media
Nivel de Limpieza:	Alta
Nivel de Concurrencia:	Media
Horario de Atención:	Todo el día. Temporada Playera: Meses de Noviembre a Abril
Ciudad:	<b>Atacames</b>
Provincia:	<b>Esmeraldas</b>



(1 votos)

**Ver  
calificaciones**

**Calificar  
Sitio**

Las diferentes actividades que usted puede practicar en **Playa de Atacames** son:

Actividades	Categoría
Bañarse en el mar	Recreación
Comprar artesanías	Otras Actividades
Descansar en hamacas	Recreación
Ir a una peña	Diversión Nocturna
Ir de pesca	Aventura
Jugar volley	Deportes
Nadar	Deportes
Navegar en botes	Aventura
Pasear en bananas	Aventura
Pasear por el malecón	Recreación
Surfear	Deportes Extremos
Tomar el sol	Recreación

Adicionalmente, si da click sobre el nombre de la ciudad, obtendrá una breve descripción de la ciudad seleccionada y la lista de todos los atractivos turísticos que desde ella puede visitar.




## Datos de la ciudad y Atractivos

Ciudad: **Guayaquil**  
 Descripción: Es capital de la provincia del Guayas, y por su gran desarrollo es la ciudad más importante del país, la más poblada y la que más riqueza genera.  
 Provincia: **Guayas**



Los atractivos que usted puede encontrar en **Guayaquil** son los siguientes:

Atractivos	Calificación	Descripción	Imagen
<b>Barrio Las Peñas</b>	☆☆☆☆☆	El barrio Las Peñas es el más antiguo de la ciudad y esta lleno de historia, y en la actualidad ha sido reparado para lucirse como una vez fue. El barrio Las Peñas es uno de los pocos legados históricos que posee la ciudad y su estilo arquitectónico muy particular, su estrecha calle, sus casas junto al río que poseen dos frentes lo convierten en un llamativo destino de visita turística. Cada mes de julio se convierte en vitrina de arte y cultura de artistas que viven en el lugar y de manifestaciones culturales que tienen la oportunidad de mostrarse año tras año.	
<b>Centro Cívico</b>	☆☆☆☆☆	Este moderno centro posee una oficina de información para turistas, salas de exposiciones, salas para convenciones, y otras amplias áreas para eventos musicales, conciertos, teatro, y otros eventos culturales. En el piso superior del centro cívico pueden ser admiradas esculturas destacadas del maestro Oswaldo Guayasamín, considerado uno de los artistas más sobresalientes del mundo entero. Este centro además posee una laguna artificial, rodeada de exuberante vegetación.	
<b>Cerro Santa Ana</b>	☆☆☆☆☆	El Cerro, llamado antiguamente Cerrito Verde, es el sitio donde se originó la ciudad, ya que en sus faldas se produjo su definitiva fundación en 1547. En la actualidad, el Cerro Santa Ana es un punto de interés turístico de la ciudad. En un recorrido de 310 metros se encuentran restaurantes, cafés, galerías de arte, cibercafés y tiendas de artesanías. Está dotado de plazoletas además de áreas verdes para la recreación y el descanso, y su más importante atractivo es la vista de la majestuosa Guayaquil. Son 456 escalones con atractivos restaurantes, gents, servicios, historia y mucho más; cada paso ascendente o descendente, en un ambiente seguro, lo pondrá en contacto con la magia del lugar.	
<b>La Rotonda</b>	☆☆☆☆☆	Es un monumento recordatorio de la célebre entrevista que tuvieron en Guayaquil, el 26 de julio de 1822, los generales Bolívar y José de San Martín. Los bronceos fueron modelados y vaciados en la ciudad de Barcelona, mientras que los altos relieves de las placas fueron fundidos en Florencia, el 25 de julio de 1937. En 1929, tanto el hemicidio como todos los trabajos en mármol, encargados íntegramente al escultor Juan Rovira, fueron debidamente terminados y entregados a la ciudad.	
<b>Malecón de El Salado</b>	☆☆☆☆☆	Es una magna obra que comprende 12 hectáreas, donde se desarrolla un complejo turístico comercial a orillas del brazo de mar conocido como Estero Salado y rodeado de maravillosos jardines tropicales. Con esta obra y a través del Boulevard 9 de Octubre, eje principal de la ciudad de este a oeste, se unen los 2 grandes proyectos de malecones de río y de brazo de mar, permitiendo al turista una visita a pie que le permitirá apreciar los diversos monumentos, plazas, parques y la arquitectura e iglesias del lugar.	
<b>Malecón Simón Bolívar (Malecón 2000)</b>	☆☆☆☆☆	Malecón 2000 es un proyecto de regeneración urbana de casi 3 Km. de extensión en donde se puede disfrutar de los grandes monumentos de la historia de Guayaquil, museos, jardines, fuentes, centros comerciales, restaurantes, bares, patios de comida, cines, así como muelles, desde donde se puede abordar embarcaciones para realizar paseos diurnos y nocturnos por el río Guayas y miradores entre otras atracciones. Constituye una de las más grandes obras emprendidas por Guayaquil y considerada modelo a nivel mundial.	
<b>Parque Histórico</b>	☆☆☆☆☆	Es una replica del Guayaquil antiguo, con frente al río, donde las personas pueden adentrarse en medio de un bosque de manglar y caminar por un sendero de madera que lo atraviesa. Busca recrear la forma de vida de años ya pasados con tres zonas: de vida silvestre, urbano arquitectónica y de exposición de tradiciones. Este sitio de 8 hectáreas está rodeado de bosques de manglares y otras especies de árboles, en el manglar existen 28 especies de animales en cautiverio, entre las que se pueden observar al venado de cola blanca, al mapache y osito lavador, grupos de saínos, tigrillos, osos perezosos, cocodrilos, etc. Existen 90 especies de aves como papagayos, pericos, águila arpía, entre otras.	
<b>Zoológico El Pantanal</b>	☆☆☆☆☆	Este zoológico es un centro de rescate y conservación de nuestra fauna donde podrá disfrutar del color y vida de la naturaleza y del placer de observar cientos de animales, en su mayoría propios del Ecuador, en un ambiente lleno de verdor y paz.	

Luego de haber obtenido información de la ciudad y de los atractivos asociados a la misma, el usuario podrá conocer todas las ciudades de la provincia a la que pertenece la ciudad seleccionada, dando click en el nombre de la provincia, como lo muestra la siguiente figura.






## Datos de la ciudad y Atractivos

### Descripción de la provincia:

Provincia **Guayas**

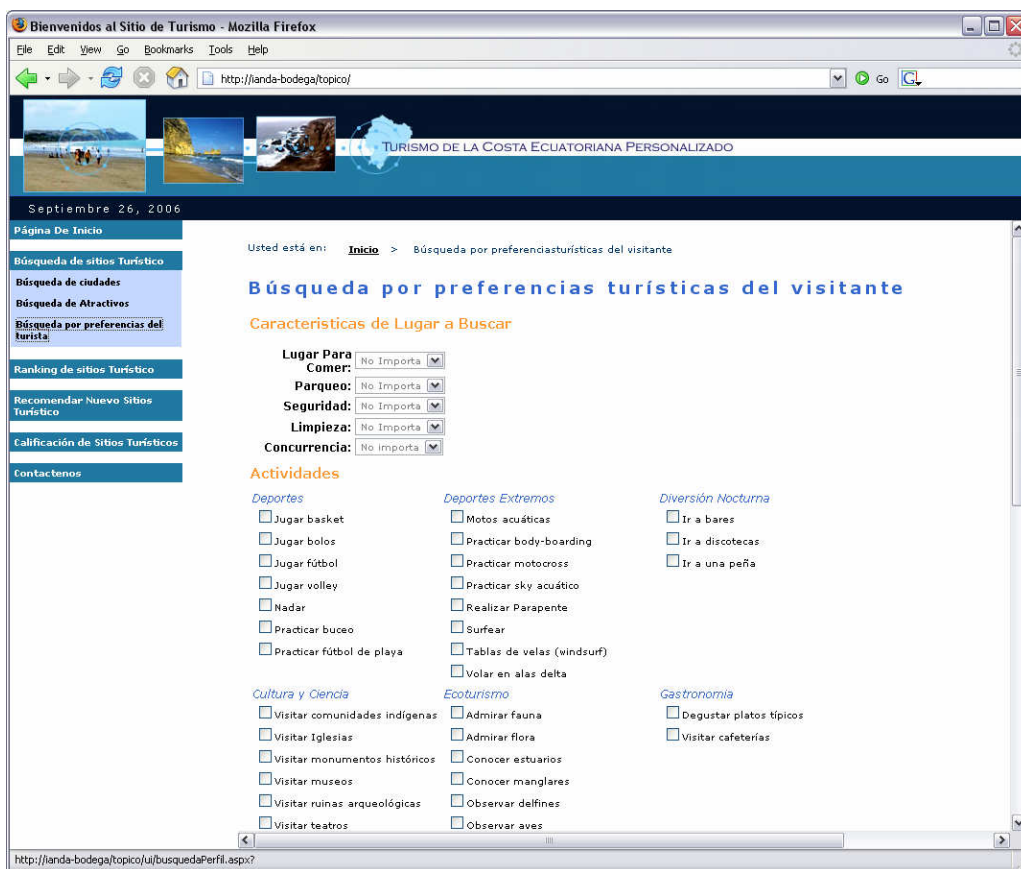
Descripción: Con amplia actividad comercial, la provincia de Guayas, ubicada al suroccidente de Ecuador, alberga al puerto principal del país, Guayaquil, así como reservas ecológicas de importancia como Manglares-Churute. Posee balnearios con buena capacidad hotelera y pequeñas playas para la práctica de algunos deportes. Tiene un aeropuerto internacional e infraestructura para la realización de eventos nacionales e internacionales. Varias cadenas hoteleras internacionales tienen filiales en Guayaquil. Su temperatura bordea los 25 grados centígrados y tiene un clima tropical húmedo.

Las ciudades que usted puede encontrar en **Guayas** son las siguientes:

Ciudad	Descripción	Imagen
<b>Ayangue</b>	Si usted desea visitar una bahía con aguas extremadamente tranquilas, nuestra sugerencia es que tome el desvío hacia Ayangue en el Km. 39. Ayangue es un pueblo dedicado a la pesca, si a usted le apasiona el buceo debería tomar un bote y dirigirse al Islote El Pelado, eso sí, vaya acompañado de alguien que conozca el sector.	
<b>Ballenita</b>	Ballenita es una tranquila ensenada, en donde se deja ver una estupenda playa convertida en balneario por propios y extraños.	
<b>Balzar</b>	Balzar es una tierra privilegiada que se considera como una zona eminentemente agrícola, ya que la generosidad de la naturaleza lo ha dotado de tierras muy fértiles.	
<b>General Villamil</b>	Es el balneario más cercano a Guayaquil. Sin lugar a dudas los mejores atractivos lo constituyen sus hermosas y extensas playas que cuenta con una excelente infraestructura hotelera y de servicios que hacen de lo más agradable la visita y estadía de miles de turistas nacionales y extranjeros.	
<b>Guayaquil</b>	Es capital de la provincia del Guayas, y por su gran desarrollo es la ciudad más importante del país, la más poblada y la que más riqueza genera.	
<b>La Libertad</b>	Los españoles lo llamaron La Agujereada o Hueca, por la configuración de su perfil costanero.	
<b>Manglaralto</b>	Está situada a orillas del mar, al noroeste de la provincia, y por	

## Buscar atractivo por perfil

Para realizar la búsqueda del lugar que más se ajusta a sus necesidades, debe dirigirse a “Búsqueda de Sitios Turísticos” y escoger la opción “Búsqueda por preferencias turísticas”



En esta ventana debe escoger las características junto con las actividades que desee realizar, para poder así obtener la información del lugar que debe visitar según sus preferencias.

Esto es, si desea un lugar que tenga parqueo, sea muy limpio, poco concurrido y no le interesa la seguridad ni el sitio para comer, en el que pueda practicar surf, buceo, ir a bares y a discotecas y volar en alas delta, debe escoger las características y actividades como se muestra en la siguiente figura:

## Búsqueda por preferencias turísticas del visitante

### Características de Lugar a Buscar

**Lugar Para Comer:** No Importa ▼

**Parqueo:** Si ▼

**Seguridad:** No Importa ▼

**Limpieza:** Alta ▼

**Concurrencia:** Baja ▼

### Actividades

#### Deportes

- Jugar basket
- Jugar bolos
- Jugar fútbol
- Jugar volley
- Nadar
- Practicar buceo
- Practicar fútbol de playa

#### Deportes Extremos

- Motos acuáticas
- Practicar body-boarding
- Practicar motocross
- Practicar sky acuático
- Realizar Parapente
- Surfear
- Tablas de velas (windsurf)
- Volar en alas delta

#### Diversión Nocturna

- Ir a bares
- Ir a discotecas
- Ir a una peña

#### Cultura y Ciencia

- Visitar comunidades indígenas
- Visitar Iglesias
- Visitar monumentos históricos
- Visitar museos
- Visitar ruinas arqueológicas
- Visitar teatros

#### Ecoturismo

- Admirar fauna
- Admirar flora
- Conocer estuarios
- Conocer manglares
- Observar delfines
- Observar aves


#### Gastronomía

- Degustar platos típicos
- Visitar cafeterías


Con estos datos, el sistema le indicará la ciudad que mejor se acopla a sus requerimientos, seguido por los atractivos que puede visitar desde esa ciudad, donde podrá encontrar todo lo que necesita para satisfacer sus necesidades. Adicionalmente se le indican tres ciudades más que también pueden interesarle conocer ya que se ajustan a sus requerimientos.

## Resultado de Búsqueda

La ciudad que mejor se asemeja a sus preferencias es:

<b>Puesto</b> # <b>1</b>	<b>Atacames</b>	
	<b>Descripción:</b> Espectaculares paisajes, mágicos atardeceres; además Atacames ofrece una rica y variada gastronomía y una amplia oferta deportiva. Atacames es el centro de la diversión de Esmeraldas. Posee más de 40 cabañas de rumba, cocteles y música que se extienden a lo largo de los 2km de malecón.	
	<b>Provincia</b> <b>Esmeraldas</b>	

Los atractivos que puede encontrar en Atacames son:

Atractivos	Descripción	Imagen
<b>Playa de Atacames</b>	Cuenta con una amplia infraestructura hotelera. A lo largo del malecón se ubican los principales restaurantes, con lo mejor de la comida típica de la región. Junto a la playa, numerosos bares ofrecen bebidas típicas a base de agua de coco, principalmente. También hay un paseo con artesanías, accesorios para el cuerpo y objetos trabajados en coral.	

Otras Opciones para sus preferencias son:

Puesto	Ciudad
2	<b>Canoa</b>
3	<b>Pedernales</b>
4	<b>Guayaquil</b>

Desde la ventana de la figura anterior puede obtener mayor información de los lugares turísticos existentes tanto en el nombre o imagen del atractivo, o puede revisar los atractivos que poseen las otras ciudades listadas, dando click sobre el nombre de dicha ciudad.

## Visualizar el ranking de sitios turísticos











Para conocer cuales son los atractivos con mejor calificación y las ciudades a las que pertenecen, debe ir a “Ranking de sitios turísticos” y seleccionar “Ranking de sitios con mejor calificación”.

En esta ventana se presentarán constantemente los lugares con mejores calificaciones y que posiblemente sean los más recomendados a la hora de realizar un viaje.

Usted está en: [Inicio](#) > [Ranking de Sitios Turísticos](#)

### Rankings de Sitios Turísticos

Atractivos con mejor Calificación:

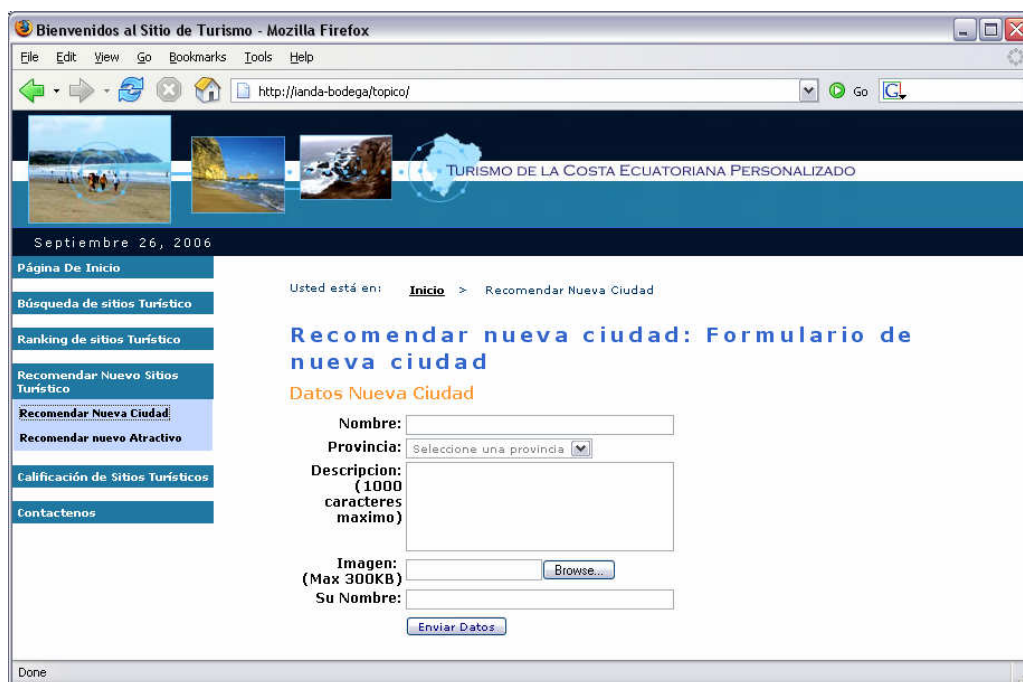
Puesto	Atractivo	Ealificación	Ciudad	Imagen
1	<a href="#">Playa de Canoa</a>	 <a href="#">Ver calificaciones</a>	Canoa	
2	<a href="#">Malecón 2000 (Malecón Simón Bolívar)</a>	 <a href="#">Ver calificaciones</a>	Guavaquil	
3	<a href="#">Playa de Montañita</a>	 <a href="#">Ver calificaciones</a>	Santa Elena	
4	<a href="#">Estación de Río Palenque</a>	 <a href="#">Ver calificaciones</a>	Palenque	
5	<a href="#">Zoológico El Pantanal</a>	 <a href="#">Ver calificaciones</a>	Guavaquil	

Done

Desde la ventana anterior podrá revisar información del atractivo, detalles de su calificación, y de la ciudad a la que pertenece respectivamente.

## Recomendar nueva ciudad

Para recomendar alguna ciudad de la costa ecuatoriana que no se haya considerado en el sistema, debe dirigirse a “Recomendar Nuevo Sitio Turístico” y elegir “Recomendar nueva ciudad”



The screenshot shows a Mozilla Firefox browser window titled "Bienvenidos al Sitio de Turismo - Mozilla Firefox". The address bar shows the URL "http://ianda-bodega/topico/". The page content includes a navigation menu on the left with options like "Página De Inicio", "Búsqueda de sitios Turístico", "Ranking de sitios Turístico", "Recomendar Nuevo Sitios Turístico", "Recomendar Nueva Ciudad", "Recomendar nuevo Atractivo", "Calificación de Sitios Turísticos", and "contactenos". The main content area displays the breadcrumb "Usted está en: Inicio > Recomendar Nueva Ciudad" and the heading "Recomendar nueva ciudad: Formulario de nueva ciudad". Below this is the "Datos Nueva Ciudad" form, which includes fields for "Nombre:", "Provincia:" (with a dropdown menu), "Descripción:" (with a limit of 1000 characters), "Imagen:" (with a "Browse..." button and a 300KB limit), and "Su Nombre:". An "Enviar Datos" button is located at the bottom of the form.

Desde esta ventana podrá ingresar los datos de una nueva ciudad que desee recomendar, para esto debe llenar un formulario que consta de:

- *Nombre:* nombre de la ciudad a recomendar.
- *Provincia:* provincia a la que pertenece la ciudad a recomendar.
- *Descripción:* breve descripción de la ciudad que está recomendando, cuales son sus principales características,
- *Imagen:* path de imagen en formato .jpg de la ciudad recomendada.
- *Su nombre:* nombre o nick de la persona que recomienda la ciudad.

## Recomendar nueva ciudad: Formulario de nueva ciudad

### Datos Nueva Ciudad

<b>Nombre:</b>	<input type="text" value="Zapotalito"/>
<b>Provincia:</b>	<input type="text" value="Los Ríos"/>
<b>Descripción:</b> (1000 caracteres maximo)	<input type="text" value="Parroquia rural del cantón Ventanas, posee una población muy antigua. Esta situado a orillas del río de su mismo nombre en la vía Ventanas-Quevedo"/>
<b>Imagen:</b> (Max 300KB)	<input type="text" value="C:\Ciudades\Zapotal.jpg"/> <input type="button" value="Browse..."/>
<b>Su Nombre:</b>	<input type="text" value="David Guzmán"/>
<input type="button" value="Enviar Datos"/>	

Una vez que haya ingresado correctamente los datos del formulario, debe presionar el botón “Enviar datos” para confirmar su recomendación. Luego se le muestra la ventana de Recomendar nuevo atractivo, para que ingrese algún atractivo que posea la ciudad recomendada.



## Recomendar nuevo atractivo

Para recomendar algún atractivo de la costa ecuatoriana que no se haya considerado en el sistema, debe dirigirse a “Recomendar Nuevo Sitio Turístico” y elegir “Recomendar nuevo atractivo”:

Usted está en: [Inicio](#) > [Recomendar Atractivo Turístico](#)

### Recomendar nuevo atractivo Turístico: Formulario Nuevo Atractivo

**Datos de ubicación del Atractivo Turístico**

Provincia:

**Datos del Atractivo**

Atractivo:

Descripción:

Dirección:

Lugar Para Comer:

Parqueo:

Nivel de Seguridad:

Nivel de Limpieza:

Nivel de Concurrencia:

Horario de Atención:

Datos Adicionales:

Imagen:

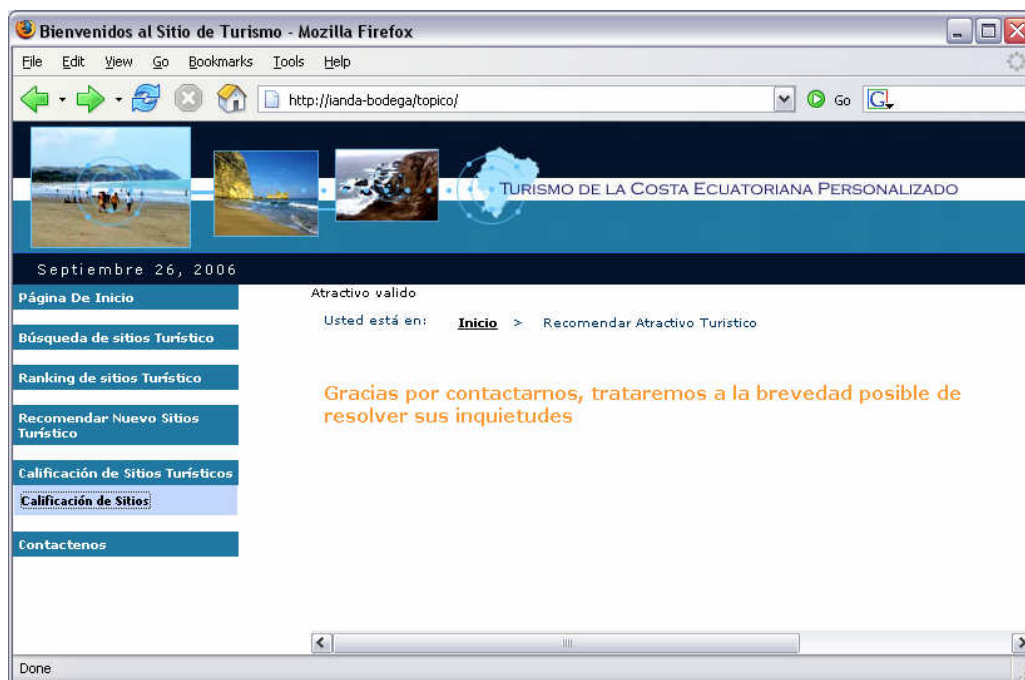
Su Nombre:

Desde la ventana de “Recomendar Nuevo Atractivo” podrá ingresar los datos de un sitio que desee dar a conocer, para esto debe llenar el formulario que consta de:

- *Provincia:* provincia a la que pertenece el atractivo.
- *Ciudad:* ciudad más cercana desde la cual se puede llegar al atractivo.
- *Atractivo:* nombre del atractivo a recomendar.
- *Descripción:* breve descripción del atractivo, sus principales características.
- *Dirección:* Ubicación geográfica del atractivo con respecto a la ciudad más cercana o a algún otro sitio de interés.
- *Lugar para comer:* debe seleccionar si o no, dependiendo de si existe o no algún lugar de comida en el atractivo.
- *Parqueo:* debe seleccionar si o no, dependiendo de si existe o no lugar de parqueo en el atractivo.
- *Nivel de seguridad:* debe seleccionar el nivel de seguridad con que cuenta el atractivo a recomendar.
- *Nivel de limpieza:* debe seleccionar el nivel de limpieza con que cuenta el atractivo a recomendar.
- *Nivel de concurrencia:* Debe seleccionar el nivel de concurrencia con que cuenta el atractivo a recomendar.
- *Horario de atención:* indicar el horario/estación/meses del año, en que el atractivo está disponible para el visitante.

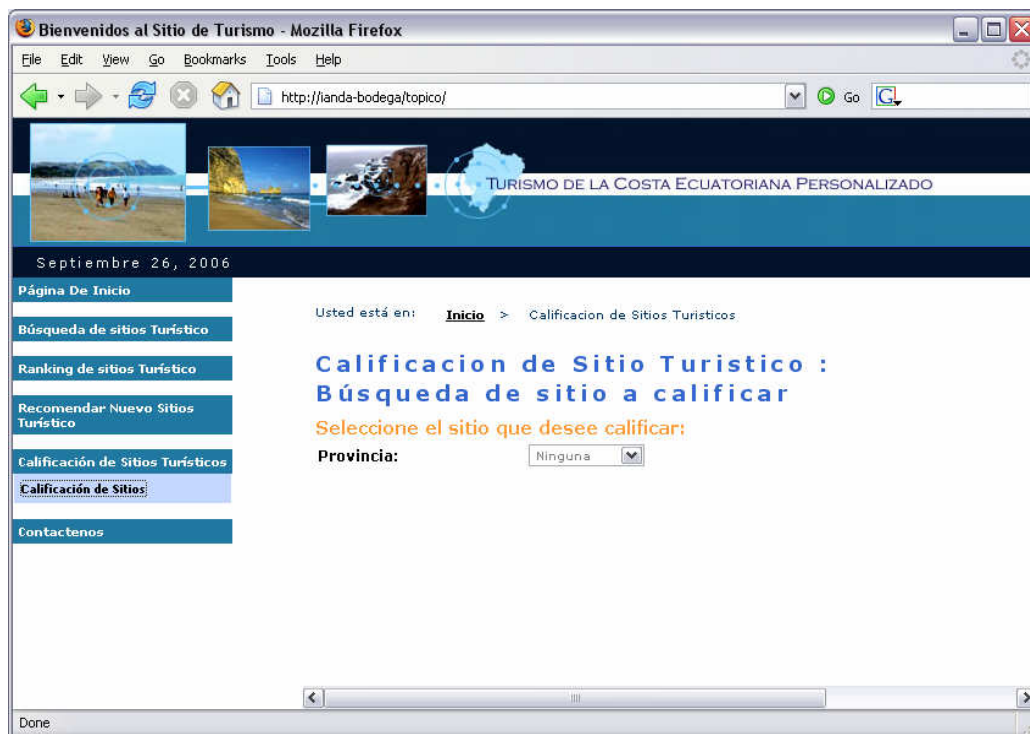
- *Datos adicionales:* en este campo puede indicar la dirección web del atractivo, el número de teléfono o algún otro campo que considere apropiado.
- *Imagen:* path de imagen adjunta del atractivo en formato .jpg.
- *Nombre:* Nombre o nick de la persona que sugiere el sitio.

Una vez que haya ingresado correctamente los datos del formulario, debe presionar el botón “Enviar datos” para confirmar su recomendación y se le mostrará un mensaje de agradecimiento por su sugerencia.



## Calificar sitio turístico

Para realizar la calificación de un sitio turístico debe dirigirse al menú “Calificación de sitios turísticos” y elegir “Calificar sitio”



Una vez que este en esta ventana deberá seleccionar el atractivo que desee calificar según la ciudad y provincia a la que pertenezca, tal como se muestra a continuación:

## Calificación de Sitio Turístico : Búsqueda de sitio a calificar

Seleccione el sitio que desee calificar:

**Provincia:** 
  
**Ciudad:** 
  
**Atractivo:**

El atractivo Seleccionado es:

**Zoológico El Pantanal**

**Descripción:** Este zoológico es un centro de rescate y conservación de nuestra fauna donde podrá disfrutar del color y vida de la naturaleza y del placer de observar cientos de animales, en su mayoría propios del Ecuador, en un ambiente lleno de verdor y paz.

**Dirección:** Esta situado en el Km 23 vía a Daule, a tan sólo 15 minutos del centro de la ciudad.

**Lugar Para Comer:** Si

**Parqueo:** No

**Nivel de Seguridad:** Alta

**Nivel de Limpieza:** Alta

**Nivel de Concurrencia:** Media

**Horario de Atención:** De lunes a domingo de 09h00 a 17h00

**Datos Adicionales:** <http://www.zoelpantanal.com>

**Ciudad:** Guayaquil

**Provincia:** Guayas



Luego de haber seleccionado estos parámetros, debe proceder a dar click en el botón “Calificar” para pasar a llenar el formulario de calificación del sitio que consta de:

- **Título:** nombre de su calificación. Ej. La mejor playa de la costa.
- **Calificación del atractivo:** nivel de calificación asignado al atractivo según su criterio. (pésimo, regular, bueno, muy bueno, excelente)
- **Seguridad:** nivel de seguridad que posee el atractivo según su criterio. (pésimo, regular, bueno, muy bueno, excelente)

- *Limpieza*: nivel de limpieza que posee el atractivo según su criterio. (pésimo, regular, bueno, muy bueno, excelente)
- *Ventajas*: Cuales son las cosas positivas que se puede decir acerca del atractivo que esta siendo calificado.
- *Desventajas*: Cuales son las cosas negativas que se puede decir acerca del atractivo que esta siendo calificado.
- *Su identificador*: nombre o nick de la persona que califica el atractivo.

## Calificación de Sitio Turístico: Formulario de calificación

El atractivo a calificar es:

Atractivo: Zoológico El Pantanal  
 Ciudad: Guayaquil  
 Provincia Guayas

Calificación

Título:	<input type="text" value="El mejor zoológico"/>
Calificación Atractivo:	<input type="text" value="Excelente"/>
Seguridad:	<input type="text" value="Muy Bueno"/>
Limpieza:	<input type="text" value="Muy Bueno"/>
Pros: (1000 caracteres máximo)	<input type="text" value="Es un lugar muy apropiado para conocer la diversidad de animales que posee nuestro país"/>
Contras: (1000 caracteres máximo)	<input type="text" value="Se encuentra un poco alejado de la ciudad"/>
Su Nombre:	<input type="text" value="David Guzmán"/>
	<input type="button" value="Calificar"/>

Luego de llenar el formulario de calificación y proceder a calificar correctamente el sitio pasará a visualizar el detalle de todas las calificaciones que tenga ese atractivo:

Septiembre 26, 2006

Usted está en: Inicio > Calificación de Sitio - Búsqueda > Calificaciones de sitio turistic

### Calificación de Sitio Turístico

**El atractivo Seleccionado:**

Atractivo: **Zoológico El Pantanal**  
 Ciudad: **Suvaqui**  
 Provincia: **Guayas**

Calificación Promedio: ☆☆☆☆ (2 votos)

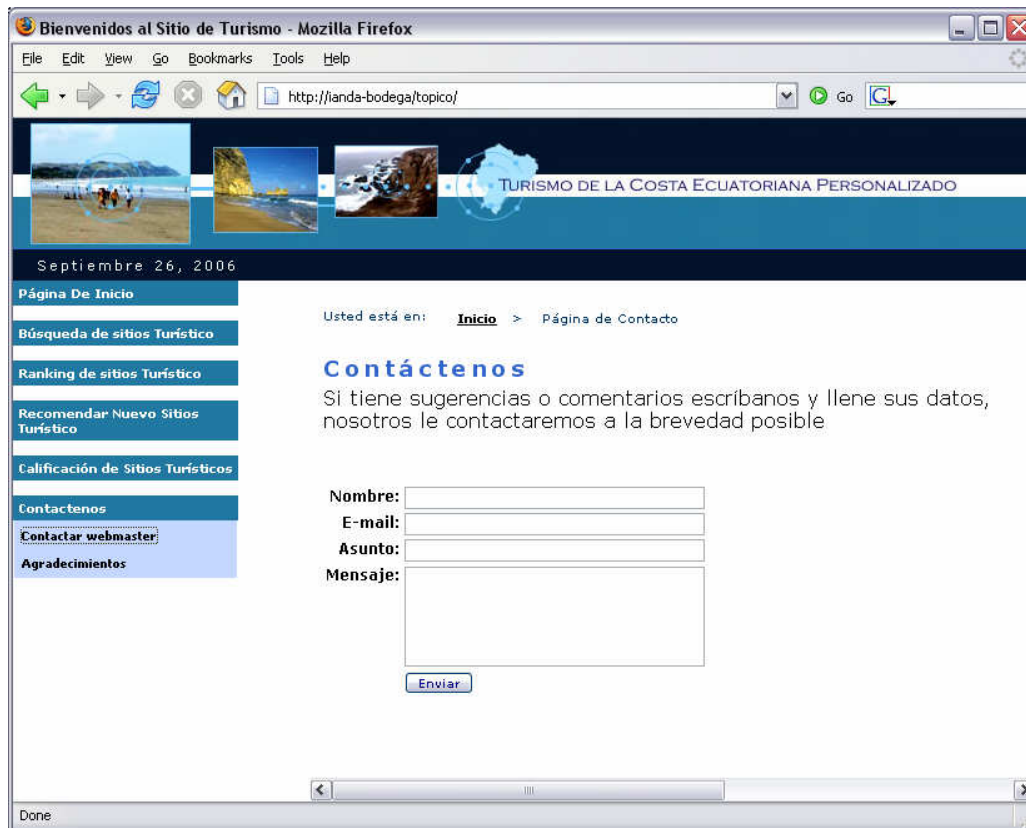
**Calificaciones**

Calificación	Seguridad	Limpieza	Título	Dejar	Comentarios	Fecha
☆☆☆☆	☆☆☆☆	☆☆☆☆	El mejor zoológico	Es un lugar muy apropiado para conocer la diversidad de animales que posee nuestro país	Se encuentra un poco alejado de la ciudad	26-09-06
☆☆☆☆	☆☆☆☆	☆☆☆☆	Excelente lugar	Este es un sitio ideal para ir a disfrutar de la naturaleza y a la vez conocer una gran variedad de animales	Debe tener cuidado al recorrer el lugar y es recomendable que siempre lo guie alguien del sitio	26-07-06

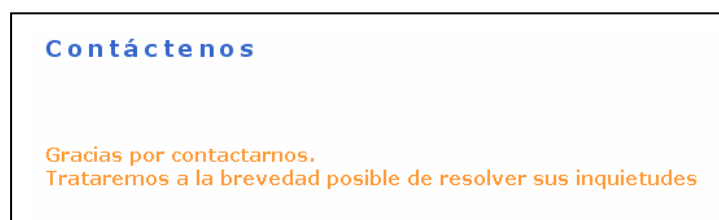
1

## Contactar al web master

Si desea realizar una sugerencia o comentarios del sistema, debe dirigirse al menú “Contáctenos” y elegir “Contactar web master”, aquí encontrará un formulario que debe llenar para comunicar sus inquietudes.



Una vez que envíe su comentario se le mostrará el siguiente mensaje de agradecimiento.





## **ANEXO 9**

### **DICCIONARIO DE DATOS DEL MODELO RELACIONAL**

#### **st\_provincia**

Entidad que guarda y almacena toda la información sobre cada una de las cinco provincias de la costa ecuatoriana.

Sus atributos son:

idProvincia.- Identificador único que se utiliza para identificar a cada una de las provincias de la costa.

Nombre.- cadena de caracteres que representa el nombre de la provincia.

Descripción.- cadena de caracteres que hace una breve descripción de la provincia y sus características más importantes dentro de esta.

#### **st\_ciudad**

Entidad que guarda información acerca de la ciudad donde se encuentra el atractivo.

Sus atributos son:

idCiudad.- identificador único que se utiliza para distinguir las diferentes ciudades que están dentro de una provincia.

Nombre.- cadena de caracteres que representa el nombre de la ciudad.

Descripción.- cadena de caracteres que describe características importantes de la ciudad.

Imagen.- cadena de caracteres que representa la dirección física de la imagen de la ciudad guardada en el servidor.

### **st\_atractivo**

Entidad que guarda información acerca de los atractivos turísticos que se encuentran en cada una de las ciudades dentro de una provincia de la costa ecuatoriana.

Sus atributos son:

idAtractivo.- Identificador único utilizado para distinguir cada uno de los atractivos.

Nombre.- cadena de caracteres que representa el nombre del atractivo.

Descripción.- cadena de caracteres que describe características importantes del atractivo.

Dirección.- cadena de caracteres que representa la dirección donde se encuentra ubicado el atractivo

Imagen.- cadena de caracteres que representa la dirección física de la imagen del atractivo guardada en el servidor.

LugarComer.- cadena de caracteres que identifica si existe o no, un lugar para comer cerca o dentro del atractivo.

Parqueo.- cadena de caracteres que identifica si tiene o no, parqueo cerca o dentro del atractivo.

HorarioAtención.- Indica el horario de atención al público en el atractivo.

Seguridad.- Integer que indica niveles de seguridad dentro del atractivo.

Es decir: baja, media y alta.

Limpio.- Integer que indica niveles de cuan limpio es el atractivo. Es decir: baja, media y alta.

Concurrencia.- Integer que indica los niveles de concurrencia en un atractivo. Estas son: alta, media y baja.

Status.- caracter que contiene el status del atractivo. Puede ser: activo o inactivo.

datosAdicionales.- cadena de caracteres que contiene información adicional del atractivo como: página web, teléfono, etc.

### **st\_tipoAtractivo**

Entidad que guarda información de los diferentes tipos de atractivos turísticos. Como son: campo, balneario, cultural, parques y malecones, comerciales y reservas ecológicas.

Sus atributos son:

idTipoAtractivo.- identificador único utilizado para determinar e identificar cada uno de los tipos de atractivos.

nombre.- cadena de caracteres que representa el nombre del tipo de atractivo.

descripción.- cadena de caracteres que representa características importantes del tipo de atractivo.

### **st\_actividad**

En esta entidad se almacena la información de las actividades que se realizan dentro de un atractivo.

Sus atributos son:

idActividad.- identificador único utilizado para identificar la actividad a realizarse dentro del atractivo.

nombre.- cadena de caracteres que representa el nombre de la actividad.

descripción.- cadena de caracteres que describen la actividad a realizarse dentro del atractivo.

**st\_categoria**

Entidad que almacena información relacionada a las categorías de las actividades que se realizan dentro de un atractivo. Categorizadas de la siguiente manera: deportes, deportes extremos, diversión nocturna, cultura y ciencia, ecoturismo, gastronomía, aventura, recreación y otras actividades.

Sus atributos son:

idCategoría.- identificador único que permite identificar la categoría de las actividades.

Nombre.- cadena de caracteres que representa el nombre de la categoría a la que pertenece la actividad.

Descripción.- cadena de caracteres para describir la categoría a la que pertenece una actividad.

**st\_revisión**

Entidad que almacena información de la calificación de los atractivos.

Sus atributos son:

idRevisión.- identificador único que permite distinguir cada una de las revisiones realizadas de los atractivos.

Descripción.- cadena de caracteres que describe características de una revisión.

Calificación.- Integer que representa el nivel de calificación del atractivo.

Pros.- cadena de caracteres que representa las ventajas del atractivo o sitio a calificar.

Cont.- cadena de caracteres que representa las desventajas o contras al calificar un sitio o atractivo.

Limpieza.- Integer que identifica los niveles de limpieza para calificar el atractivo.

Seguridad.- Integer que identifica los niveles de seguridad para calificar el atractivo.

Posteador.- cadena de caracteres que representa el nombre de la persona que califica el atractivo.

FechaCreación.- cadena de caracteres que representa día, mes y año en que fue creada o ingresada la calificación del atractivo.

## **DICCIONARIO DE DATOS DEL DISEÑO MULTIDIMENSIONAL**

### **md\_provincia**

Entidad que guarda y almacena toda la información sobre cada una de las cinco provincias de la costa ecuatoriana.

Sus atributos son:

idProvincia.- Identificador único que se utiliza para identificar a cada una de las provincias de la costa.

Nombre.- cadena de caracteres que representa el nombre de la provincia.

Descripción.- cadena de caracteres que hace una breve descripción de la provincia y sus características más importantes dentro de esta.

### **md\_ciudad**

Entidad que guarda información acerca de la ciudad donde se encuentra el atractivo.

Sus atributos son:

idCiudad.- identificador único que se utiliza para distinguir las diferentes ciudades que están dentro de una provincia.

Nombre.- cadena de caracteres que representa el nombre de la ciudad.

### **md\_atractivo**

Entidad que guarda información acerca de los atractivos turísticos que se encuentran en cada una de las ciudades dentro de una provincia de la costa ecuatoriana.

Sus atributos son:

idAtractivo.- Identificador único utilizado para distinguir cada uno de los atractivos.

Nombre.- cadena de caracteres que representa el nombre del atractivo.

Descripción.- cadena de caracteres que describe características importantes del atractivo.

Dirección.- cadena de caracteres que representa la dirección donde se encuentra ubicado el atractivo

LugarComer.- cadena de caracteres que identifica si existe o no, un lugar para comer cerca o dentro del atractivo.

Parqueo.- cadena de caracteres que identifica si tiene o no, parqueo cerca o dentro del atractivo.

Seguridad.- Integer que indica niveles de seguridad dentro del atractivo.

Es decir: baja, media y alta.

Limpio.- Integer que indica niveles de cuan limpio es el atractivo. Es decir: baja, media y alta.

Concurrencia.- Integer que indica los niveles de concurrencia en un atractivo. Estas son: alta, media y baja.



**md\_tipoAtractivo**

Entidad que guarda información de los diferentes tipos de atractivos turísticos. Como son: campo, balneario, cultural, parques y malecones, comerciales y reservas ecológicas.

Sus atributos son:

idTipoAtractivo.- identificador único utilizado para determinar e identificar cada uno de los tipos de atractivos.

nombre.- cadena de caracteres que representa el nombre del tipo de atractivo.

descripción.- cadena de caracteres que representa características importantes del tipo de atractivo.

**md\_actividad**

En esta entidad se almacena la información de las actividades que se realizan dentro de un atractivo.

Sus atributos son:

idActividad.- identificador único utilizado para identificar la actividad a realizarse dentro del atractivo.

nombre.- cadena de caracteres que representa el nombre de la actividad.

descripción.- cadena de caracteres que describen la actividad a realizarse dentro del atractivo.

### **md\_categoria**

Entidad que almacena información relacionada a las categorías de las actividades que se realizan dentro de un atractivo. Categorizadas de la siguiente manera: deportes, deportes extremos, diversión nocturna, cultura y ciencia, ecoturismo, gastronomía, aventura, recreación y otras actividades.

Sus atributos son:

idCategoría.- identificador único que permite identificar la categoría de las actividades.

Nombre.- cadena de caracteres que representa el nombre de la categoría a la que pertenece la actividad.

Descripción.- cadena de caracteres para describir la categoría a la que pertenece una actividad.

### **md\_año**

Entidad que almacena información del año como parte de la unidad de tiempo para así dimensionar el modelo multidimensional de nuestro sistema.

Sus atributos son:

idAnio.- Identificador único que permite identificar el año

anio.- Integer que describe el año como referencia.

### **md\_mes**

Entidad que almacena información del mes como unidad de tiempo para la multidimensionalidad.

Sus atributos son:

id.- Identificador único que permite identificar el mes.

mes.- Integer que identifica el mes como un valor de fecha.

### **md\_semana**

Entidad que almacena información de la semana como unidad de tiempo para dimensionar el modelo multidimensional de nuestro sistema.

Sus atributos son:

id.- identificador único que permite identificar la semana.

semana.- cadena de caracteres que identifica la semana como un valor de fecha.

fecha.- De tipo fecha identifica semana, mes y año como una unidad de tiempo para saber cuales han sido los valores que toma en ese periodo cierto atractivo o ciudad.

## **TABLA DE HECHOS EN EL DISEÑO MULTIDIMENSIONAL**

### **md\_atractivoCalificación**

Entidad o tabla de hecho que guarda cada uno de los puntajes y calificaciones del atractivo.

Sus atributos son:

id.- Identificador único que permite identificar la calificación del atractivo.

puntajeDeportes.- Decimal que indica el puntaje obtenido en las actividades de deportes.

puntajeDepEx.- Decimal que indica el puntaje obtenido en las actividades de Deportes Extremos.

puntajeDivNoct.- Decimal que indica el puntaje obtenido en las actividades de Diversión Nocturna.

puntajeCultura.- Decimal que indica el puntaje obtenido en las actividades de Cultura.

puntajeEcoturismo.- Decimal que indica el puntaje obtenido en las actividades de Ecoturismo

puntajeAventura.- Decimal que indica el puntaje obtenido en las actividades de Aventura.

puntajeRecreación.- Decimal que indica el puntaje obtenido en las actividades de Recreación.

calificaciónseg.- Valor decimal que indica la calificación promedio dada al atractivo por su seguridad.

calificaciónConcurrencia.- Valor decimal que indica la calificación promedio dada al atractivo por su concurrencia.

calificaciónAtractivo.- Valor decimal que indica la calificación promedio dada al atractivo mismo como lugar.

### **md\_ciudadCalificación**

Entidad o tabla de hecho que guarda cada uno de los puntajes y calificaciones de una ciudad.

Sus atributos son:

id.- Identificador único que permite identificar la calificación de la ciudad.

calificaciónSeguridad.- Valor decimal que indica la calificación promedio dada a la ciudad por su seguridad.

calificaciónLimpieza.- Valor decimal que indica la calificación promedio dada a la ciudad por su limpieza.

calificaciónCiudad.- Valor decimal que indica la calificación promedio dada a la ciudad misma como lugar.

puntajeCampo.- Decimal que indica el puntaje obtenido de la ciudad como campo.

puntajeCultura.- Decimal que indica el puntaje obtenido de la ciudad en base a la cultura.

puntajeMonumentos.- Decimal que indica el puntaje obtenido de la ciudad por sus monumentos.

puntajeComercio.- Decimal que indica el puntaje obtenido de la ciudad por el comercio.

puntajeReservas.- Decimal que indica el puntaje obtenido de la ciudad por sus reservas ecológicas.

ranking.- Integer que indica el puesto que obtuvo la ciudad o atractivo.

parqueoProb.- Decimal que indica la probabilidad de que en la ciudad exista parqueos.

lugarComidaProb.- Decimal que indica la probabilidad de que en la ciudad exista un lugar para comer.

concuranciaProb.- Decimal que indica la probabilidad de que en la ciudad haya mucha concurrencia de visitantes.

## ANEXO 10

### ANÁLISIS DE ALGORITMOS A IMPLEMENTARSE

Para la realización de las pruebas previas para determinar cuál era el algoritmo más adecuado para el desarrollo del sistema, se escogió una pequeña muestra de datos, como se muestra en la siguiente tabla. Con la finalidad de explicar adecuadamente el funcionamiento de cada método y junto con sus ventajas y desventajas.

ATRIBUTOS				CLASE
Bañarse en el mar	Observar aves	Visitar ruinas arqueológicas	Visitar museos	Tipo de Atractivo
1	1	1	1	3
0	0	0	1	3
0	0	0	0	1
1	0	0	0	4
0	0	0	0	3
0	0	1	1	4
0	1	1	1	4
0	0	1	0	1
0	0	0	0	3

Esta tabla posee sus elementos discretizados de tal manera que el valor numérico “1” corresponde a “Si” y el valor “0” corresponde a “No”

## ANEXO 10A

### EJEMPLO DE APLICACIÓN DE ALGORITMO ID3

Una vez dada la matriz de la tabla anterior, debemos calcular la ganancia de cada atributo para seleccionar aquel que tenga mayor ganancia de información. La ganancia viene dada por:

$$G(A) = E(C) - E(A) ,$$

donde:

$G(A)$  = Ganancia del atributo

$E(C)$  = Entropía de la clase

$E(A)$  = Entropía del atributo

Las entropías son calculadas usando las siguientes fórmulas:

$$E = \sum_{i=1}^C - p_i \log_2 p_i$$

En nuestro ejemplo tenemos tres clases (1, 3, 4), y calculamos la entropía de la clase de la siguiente forma:

$$E(C) = \sum_{i=1}^C - p_i \log_2 p_i = -\frac{2}{9} \log_2 \left(\frac{2}{9}\right) - \frac{4}{9} \log_2 \left(\frac{4}{9}\right) - \frac{3}{9} \log_2 \left(\frac{3}{9}\right)$$

CLASE	1	3	4	Total
	2	4	3	9
Entropía	1,53049			



Para determinar la entropía del atributo “Bañarse en el mar” realizamos el siguiente procedimiento:

$$E(C_{B\_mar=1}) = 0 - \frac{1}{2} \log_2\left(\frac{1}{2}\right) - \frac{1}{2} \log_2\left(\frac{1}{2}\right) = 1$$

$$E(C_{B\_mar=0}) = -\frac{2}{7} \log_2\left(\frac{2}{7}\right) - \frac{3}{7} \log_2\left(\frac{3}{7}\right) - \frac{2}{7} \log_2\left(\frac{2}{7}\right) = 1,556656$$

$$E(C_{B\_mar}) = \frac{2}{9} \times 1 + \frac{7}{9} \times 1,556656 = 1,43295$$

$$G(B\_mar) = 1,53049 - 1,43295 = 0,09754$$

Estos resultados, junto con los resultados de la primera iteración del algoritmo los se muestran en las siguientes tablas:

#### Determinar el nodo raíz

Bañarse mar	Si	No
1	0	2
3	1	3
4	1	2
<b>Total</b>	2	7
<b>Entropía</b>	1	1,55665
<b>E(Bañarse mar)</b>		1,43295
<b>Ganancia</b>		0,09754

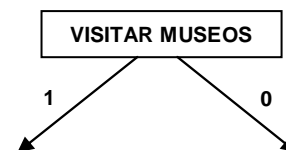
Observar Aves	Si	No
1	0	2
3	1	3
4	1	2
<b>Total</b>	2	7
<b>Entropía</b>	1	1,556656
<b>E(Observar Aves)</b>		1,432955
<b>Ganancia</b>		0,09754

Visitar Ruinas	Si	No
1	1	1
3	1	3
4	2	1
<b>Total</b>	4	5
<b>Entropía</b>	1,5	1,370950
<b>E(Visitar Ruinas)</b>		1,428305
<b>Ganancia</b>		0,10219

Visitar Museos	Si	No
1	0	2
3	2	2
4	2	1
<b>Total</b>	4	5
<b>Entropía</b>	1	1,521928
<b>E(Visitar Museos)</b>		1,289960
<b>Ganancia</b>		0,24053

Una vez que hemos obtenido la ganancia de cada atributo, escogemos aquel que tienen mayor ganancia como nodo raíz del árbol y pasamos a la siguiente iteración.

Ganancias	
Bañarse en mar	0,097538
Observar Aves	0,097538
Visitar Ruinas	0,102187
Visitar Museos	<b>0,240533</b>



**Visitar Museos = 1**

Continuando con el proceso vamos a determinar que ocurre el atributo Visitar Museos toma el valor 1.

De la matriz de datos se toman aquellos valores donde el atributo Visitar Museos toma el valor 1 y se elimina la columna correspondiente a dicho atributo para volver a realizar los cálculos anteriores con la matriz resultante, como se muestra a continuación:

Bañarse en el mar	Observar aves	Visitar ruinas arqueológicas	Visitar museos	Tipo de Atractivo
1	1	1	1	3
0	0	0	1	3
0	0	0	0	1
1	0	0	0	4
0	0	0	0	3
0	0	1	1	4
0	1	1	1	4
0	0	1	0	1
0	0	0	0	3

Matriz Inicial

Bañarse en el mar	Observar aves	Visitar ruinas arqueológicas	Tipo de Atractivo
1	1	1	3
0	0	0	3
0	0	1	4
0	1	1	4

Matriz Resultante

CLASE	1	3	4	Total
	0	2	2	4
Entropía	1,00000			

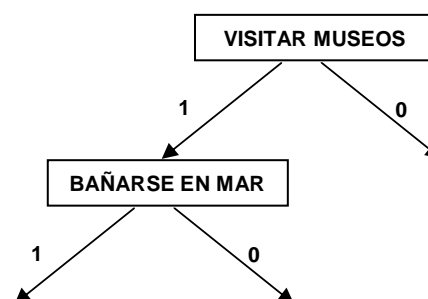
Bañarse en mar	Si	No
3	1	1
4	0	2
<b>Total</b>	1	3
<b>Entropía</b>	0	0,918295
<b>E(Bañarse_mar)</b>		0,688721
<b>Ganancia</b>		0,31128

Observar Aves	Si	No
3	1	1
4	1	1
<b>Total</b>	2	2
<b>Entropía</b>	1	1
<b>E(Observar Aves)</b>		1
<b>Ganancia</b>		0,000

Visitar Ruinas	Si	No
3	1	1
4	2	0
<b>Total</b>	3	1
<b>Entropía</b>	0,918295	0
<b>E(Visitar ruinas)</b>		0,688721
<b>Ganancia</b>		0,31128

Como se puede apreciar en la siguiente tabla, las ganancias de atributos tienen el mismo valor, lo cual indica que ambos tienen igual incidencia en los resultados, por lo que es indiferente el atributo escogido como siguiente nodo del árbol.

Ganancias	
Bañarse en mar	0,311278
Observar Aves	0
Visitar Ruinas	0,311278

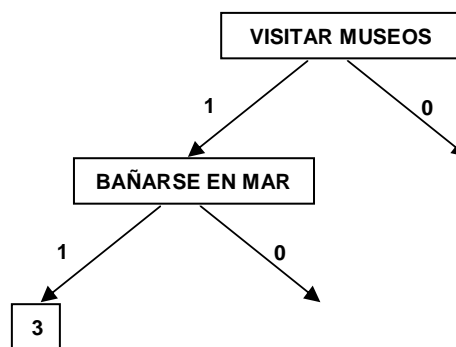


**Bañarse en mar = 1**

Al obtener la matriz resultante considerando que el atributo Bañarse en el mar tome el valor 1, se determina que dicho atributo pertenece al valor de la clase 3.

Bañarse en el mar	Observar aves	Visitar ruinas arqueológicas	Tipo de Atractivo
1	1	1	3

Matriz Resultante

**Bañarse en mar = 0**

Observar aves	Visitar ruinas arqueológicas	Tipo
0	0	3
0	1	4
1	1	4

Matriz Resultante

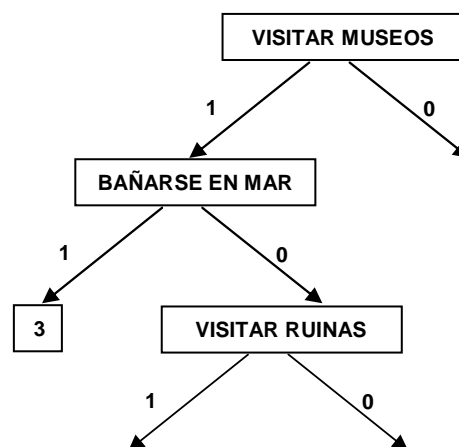
CLASE	1	3	4	Total
	0	1	2	3

Observar Aves	Si	No
3	0	1
4	1	1
<b>Total</b>	1	2
<b>Entropía</b>	0	1
<b>E(Observar Aves)</b>		0,5
<b>Ganancia</b>		0,50000

Visitar Ruinas	Si	No
3	0	1
4	2	0
<b>Total</b>	2	1
<b>Entropía</b>	0	0
<b>E(Visitar Ruinas)</b>		0
<b>Ganancia</b>		1,00000

Como se puede observar en esta iteración, cuando el atributo “Bañarse en le mar” toma el valor “0”, se tiene que el atributo con mayor ganancia es “Visitar Ruinas”, generando el siguiente árbol:

Ganancias	
Observar Aves	0,5
Visitar Ruinas	1



### Visitar Ruinas = 1

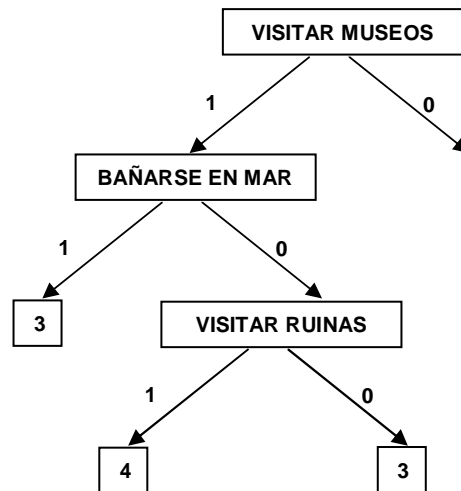
Se puede notar claramente que el atributo Visitar Ruinas nos lleva a las hojas del árbol, cuando Visitar Ruinas toma el valor de 1, la clase toma el valor “4”, como se ve en la siguiente tabla:

Observar aves	Visitar ruinas arqueológicas	Tipo de Atractivo
0	0	3
0	1	4
1	1	4

### Visitar Ruinas = 0

Mientras que, cuando Visitar ruinas toma el valor de “0”, la clase Tipo de Atractivo adquiere el valor “3”, llegando a las hojas del árbol, como se puede ver a continuación:

Observar aves	Visitar ruinas arqueológicas	Tipo de Atractivo
0	0	3
0	1	4
1	1	4



### Visitar Museos = 0

Continuando hacia la derecha del árbol, se realiza el análisis para Visitar Museos = 0, obteniendo los siguientes resultados:

Bañarse en el mar	Observar aves	Visitar ruinas arqueológicas	Tipo de Atractivo
0	0	0	1
1	0	0	4
0	0	0	3
0	0	1	1
0	0	0	3

CLASE	1	3	4	Total
	2	2	1	5
Entropía	1,52193			



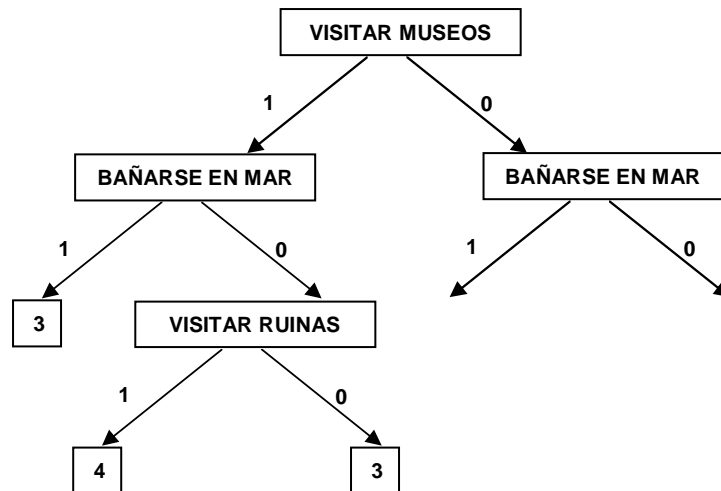
Bañarse en mar	Si	No
1	0	2
3	0	2
4	1	0
<b>Total</b>	1	4
<b>Entropía</b>	0	1
<b>E(Bañarse mar)</b>		0,8
<b>Ganancia</b>		0,72193

Observar Aves	Si	No
1	0	2
3	0	2
4	0	1
<b>Total</b>	0	5
<b>Entropía</b>	0	1,521928
<b>E(Observar Aves)</b>		1,521928
<b>Ganancia</b>		0,00000

Visitar Ruinas	Si	No
1	1	1
3	0	2
4	0	1
<b>Total</b>	1	4
<b>Entropía</b>	0	1,5
<b>E(Visitar Ruinas)</b>		1,2
<b>Ganancia</b>		0,32193

En esta iteración resulta que el atributo con mayor ganancia es “Bañarse en mar”:

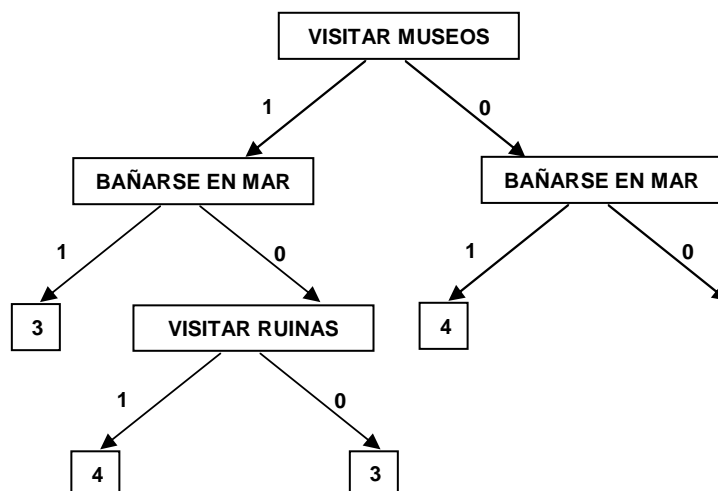
Ganancias	
Bañarse mar	0,721928
O. Aves	0
Visi Ruinas	0,321928



**Bañarse en mar = 1**

Si Bañarse en el mar toma el valor "1" se tiene que la clase (tipo de atractivo) toma el valor "4":

Bañarse en el mar	Observar aves	Visitar ruinas arqueológicas	Tipo de Atractivo
1	0	0	4



**Bañarse en mar = 0**

Cuando Bañarse en el mar toma el valor "0" se obtienen los resultados que se muestran a continuación:

Observar aves	Visitar ruinas arqueológicas	Tipo de Atractivo
0	0	1
0	0	3
0	1	1
0	0	3

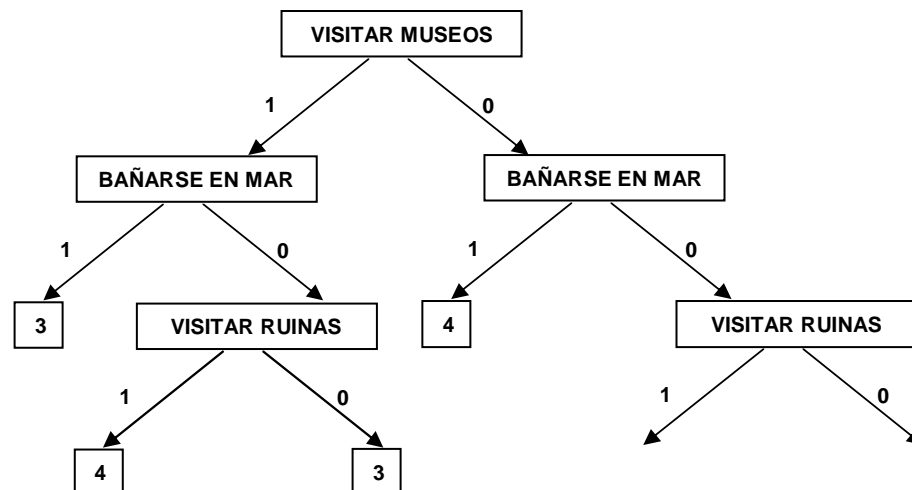
CLASE	1	3	4	Total
	2	2	0	4
Entropía	1,00000			

Observar aves	Si	No
1	0	2
3	0	2
<b>Total</b>	0	4
<b>Entropía</b>	0	1
<b>E(Observar aves)</b>		1
<b>Ganancia</b>		0,00000

Visitar Ruinas	Si	No
1	1	1
3	0	2
<b>Total</b>	1	3
<b>Entropía</b>	0	0,918295
<b>E(Visitar Ruinas)</b>		0,688721
<b>Ganancia</b>		0,31128

Podemos notar que el atributo con mayor ganancia es Visitar Ruinas:

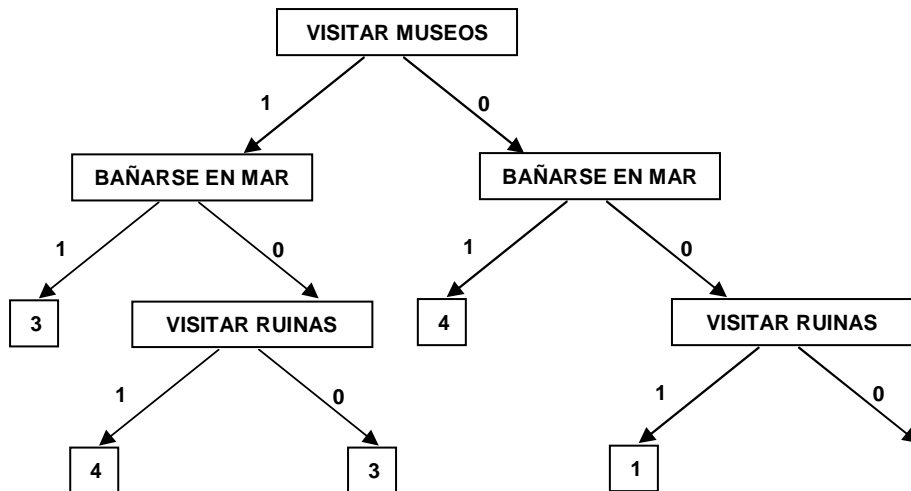
Ganancias	
Observar Aves	0
Visitar Ruinas	0,311278



**Visitar Ruinas = 1**

Si Visitar Ruinas toma el valor "1" llegamos a una hoja del árbol con Tipo de Atractivo = 1:

Observar Aves	Visitar Ruinas	Tipo de Atractivo
0	1	1



**Visitar Ruinas = 0**

Mientras que si Visitar Ruinas toma el valor "0" obtenemos los siguientes resultados:

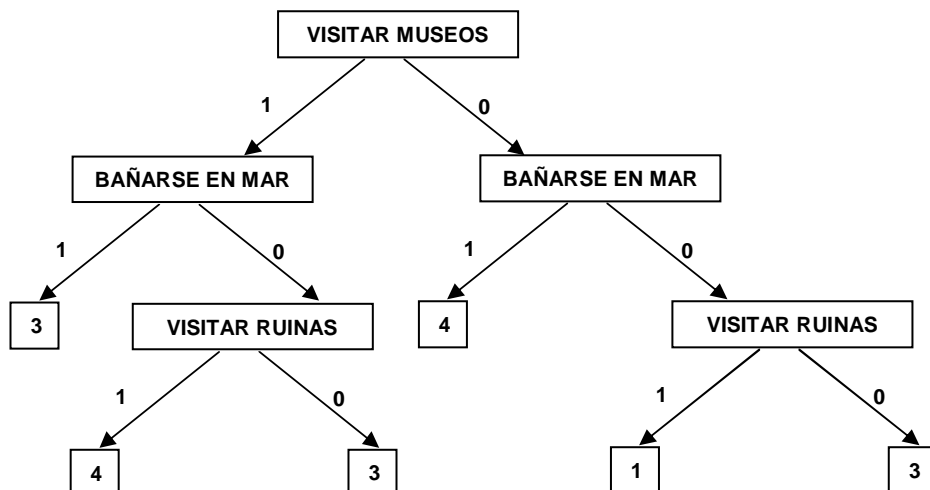
Observar aves	Visitar ruinas arqueológicas	Tipo de Atractivo
0	0	1
0	0	3
0	1	1
0	0	3

O. Aves	Tipo
0	1
0	3
0	3

<b>CLASE</b>	1	3	4	Total
	1	2	0	3
Entropía	0,91830			

Observar Aves	Si	No
1	0	1
3	0	2
<b>Total</b>	0	3
<b>Entropía</b>	0	0,918295
<b>E(Observar Aves)</b>		0,918295
<b>Ganancia</b>		0,00000

Como se puede notar la ganancia del atributo Observar Aves es igual a cero "0", por lo que tal atributo no es considerado en el árbol y se toma como hoja para este nodo el valor de la clase con mayor entropía, esto es "3", dando como resultado final el árbol que se muestra a continuación:



Con esto se demuestra que el algoritmo ID3 descarta aquellas variables cuya ganancia no aporta información al modelo, por lo que fue descartado para el desarrollo del sistema, ya que en muchas ocasiones el usuario podría requerir aquellas variables y no obtendría un resultado adecuado.

## ANEXO 10B

### EJEMPLO DE APLICACIÓN DE ALGORITMO VECINOS MÁS PRÓXIMOS

Dada la matriz del anexo 10, el algoritmo de Kmean crea grupos, según la distancia Euclídea entre atributos, para luego evaluar a que grupo pertenece cada recorrido.

Inicialmente se dividen los datos en los grupos A, B y C. como se muestra a continuación:

Grupo	Orden Inicial	Bañarse en mar	Observar aves	Visitar ruinas	Visitar museos	TIPO
A	1	1	1	1	1	3
A	2	0	0	0	1	3
A	3	0	0	0	0	1
B	4	1	0	0	0	4
B	5	0	0	0	0	3
B	6	0	0	1	1	4
C	7	0	1	1	1	4
C	8	0	0	1	0	1
C	9	0	0	0	0	3

Luego calculamos las medias de cada grupo, para cada atributo, obteniendo la tabla de centroides de cada grupo:

MEDIAS	Bañarse en el mar	Observar aves	Visitar ruinas	Visitar museos
gA	0,333	0,333	0,333	0,667
gB	0,333	0,000	0,333	0,333
gC	0,000	0,333	0,667	0,333

Aplicando la distancia Euclídea, que se muestra en la siguiente fórmula, se obtiene la distancia de cada elemento al centroide de los grupos, como se muestra en la siguiente tabla.

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

Distancia a los Grupos				
A	B	C	Distancia Mínima	Grupo
1,202	1,528	1,414	1,202	A
0,667	0,816	1,000	0,667	A
0,882	0,577	0,816	0,577	B
1,054	0,816	1,291	0,816	B
0,882	0,577	0,816	0,577	B
0,882	1,000	0,816	0,816	C
1,054	1,414	1,000	1,000	C
1,054	0,816	0,577	0,577	C
0,882	0,577	0,816	0,577	B

De entre estas distancias se encasilla cada elemento al grupo que tenga menor distancia, para generar los nuevos grupos como se muestra en la tabla a continuación:



Grupo	Orden	Bañarse en mar	Observar aves	Visitar ruinas	Visitar museos	TIPO
A	1	1	1	1	1	3
A	2	0	0	0	1	3
B	3	0	0	0	0	1
B	4	1	0	0	0	4
B	5	0	0	0	0	3
B	9	0	0	0	0	3
C	6	0	0	1	1	4
C	7	0	1	1	1	4
C	8	0	0	1	0	1

Una vez obtenida esta matriz, realizamos nuevamente el mismo procedimiento anterior para determinar si se generan nuevos cambios en los grupos.

Obtenemos nuevamente la media de cada grupo:

MEDIAS	Bañarse en el mar	Observar aves	Visitar ruinas arqueológicas	Visitar museos
gA	0,500	0,500	0,500	1,000
gB	0,250	0,000	0,000	0,000
gC	0,000	0,333	1,000	0,667

Generamos la nueva tabla de distancias:

Distancia a los Grupos				
A	B	C	Distancia Mínima	Grupo
0,866	1,887	1,247	0,866	A
0,866	1,031	1,106	0,866	A
1,323	0,250	1,247	0,250	B
1,323	0,750	1,599	0,750	B
1,323	0,250	1,247	0,250	B
1,323	0,250	1,247	0,250	B
0,866	1,436	0,471	0,471	C
0,866	1,750	0,745	0,745	C
1,323	1,031	0,745	0,745	C

En esta iteración se puede observar que no existen cambios en la clasificación de los grupos, por lo que se dice que el algoritmo converge, y los grupos resultantes son:

Grupo	Orden Final	Bañarse en mar	Observar aves	Visitar ruinas	Visitar museos	TIPO
A	1	1	1	1	1	3
A	2	0	0	0	1	3
A	3	0	0	0	0	1
B	4	1	0	0	0	4
B	5	0	0	0	0	3
B	6	0	0	1	1	4
C	7	0	1	1	1	4
C	8	0	0	1	0	1
C	9	0	0	0	0	3

Una vez realizada la clasificación en grupos de los datos, pasamos a evaluar el método con el siguiente recorrido:

*"Bañarse en el mar" = 0 y "Observar aves" = 1 y "Visitar Ruinas" = 1 Y*

*"Visitar museos" = 0*

Grupo	Orden Final	Bañarse en mar	Observar aves	Visitar ruinas	Visitar museos	TIPO
A	1	1	1	1	1	3
A	2	0	0	0	1	3
A	3	0	0	0	0	1
B	4	1	0	0	0	4
B	5	0	0	0	0	3
B	6	0	0	1	1	4
C	7	0	1	1	1	4
C	8	0	0	1	0	1
C	9	0	0	0	0	3
<b>?</b>		<b>0</b>	<b>1</b>	<b>1</b>	<b>0</b>	

Calculamos la distancia Euclídea de los nuevos datos a cada uno de los centroides de los grupos.

<b>Distancia a los Grupos</b>				
<b>A</b>	<b>B</b>	<b>C</b>	<b>Distancia Mínima</b>	<b>Grupo</b>
1,323	1,436	0,943	0,943	C

Obteniendo como resultado que el recorrido antes mencionado pertenece al grupo C.

El problema existente en la aplicación del algoritmo de los vecinos más próximos en el sistema era que su clasificación no coincidía con la clasificación previa (por tipo de atractivo) que contienen los datos, y adicionalmente, al aumentar el número de variables para la predicción, éste se volvía menos eficiente, por lo que no conducía a óptimos resultados.

## ANEXO 10C

### EJEMPLO DE APLICACIÓN DE ALGORITMO NAIVE BAYES

Para poder aplicar el algoritmo de Naive Bayes, se parte de la matriz del anexo 10, de la cual se deben generar las matrices de frecuencias y de probabilidades  $P(A_i/c)$  de cada uno de los atributos, junto con la matriz de probabilidades a priori y de frecuencia de la clase  $P(c)$ , como se muestra a continuación:

MATRIZ DE DATOS				
Bañarse en el mar	Observar aves	Visitar ruinas	Visitar museos	Tipo de Atractivo
1	1	1	1	3
0	0	0	1	3
0	0	0	0	1
1	0	0	0	4
0	0	0	0	3
0	0	1	1	4
0	1	1	1	4
0	0	1	0	1
0	0	0	0	3

Matriz de Frecuencia de la clase	
Tipo 1	2
Tipo 3	4
Tipo 4	3
TOTAL	9

Matriz de Probabilidad a priori de la Clase		
Tipo 1	2/9=	0,222222222
Tipo 3	4/9=	0,444444444
Tipo 4	3/9=	0,333333333

### Matrices de Frecuencia de cada atributo

Bañarse en el mar (BM)			
	Tipo 1	Tipo 3	Tipo 4
SI	0	1	1
NO	2	3	2

Observar Aves (OA)			
	Tipo 1	Tipo 3	Tipo 4
SI	0	1	1
NO	2	3	2

Visitar Ruinas (VR)			
	Tipo 1	Tipo 3	Tipo 4
SI	1	1	2
NO	1	3	1

Visitar Museos (VM)			
	Tipo 1	Tipo 3	Tipo 4
SI	0	2	2
NO	2	2	1

### Matrices de probabilidades de cada atributo

Para determinar las matrices de probabilidad condicional, se utilizan la matriz de frecuencias de cada atributo, y la matriz de frecuencia de la clase, usando la siguiente fórmula:

$$P(a_i / P(a_i)) = \frac{n(a_i / P(a_i))}{n(P(a_i))}$$

Donde

$n(a_i / P(a_i))$ : La frecuencia con que el atributo  $A_i$  toma el valor  $a_i$ , dada la clase  $c_j$ .

$n(P(a_i)) =$  La frecuencia de la clase  $c_j$ .

A continuación se muestran las matrices de probabilidad de cada atributo:

<b>Bañarse en el mar (BM)</b>			
<b>P(BM/Tipo)</b>	<b>P(Tipo=1)</b>	<b>P(Tipo=2)</b>	<b>P(Tipo=2)</b>
<b>P(BM=Si)</b>	0	0,25	0,33333333
<b>P(BM=No)</b>	1	0,75	0,66666667

<b>Observar Aves (OA)</b>			
<b>P(OA/Tipo)</b>	<b>P(Tipo=1)</b>	<b>P(Tipo=2)</b>	<b>P(Tipo=2)</b>
<b>P(OA=Si)</b>	0	0,25	0,33333333
<b>P(OA=No)</b>	1	0,75	0,66666667

<b>Visitar Ruinas (VR)</b>			
<b>P(VR/Tipo)</b>	<b>P(Tipo=1)</b>	<b>P(Tipo=2)</b>	<b>P(Tipo=2)</b>
<b>P(VR=Si)</b>	0,5	0,25	0,66666667
<b>P(VR=No)</b>	0,5	0,75	0,33333333

<b>Visitar Museos (VM)</b>			
<b>P(VM/Tipo)</b>	<b>P(Tipo=1)</b>	<b>P(Tipo=2)</b>	<b>P(Tipo=2)</b>
<b>P(VM=Si)</b>	0	0,5	0,66666667
<b>P(VM=No)</b>	1	0,5	0,33333333

Estas matrices ayudan a generar el vector de probabilidades asociado a cada recorrido de la red bayesiana. Por ejemplo, vamos a predecir a que tipo de atractivo corresponde el siguiente recorrido (R):

"Bañarse en el mar" = 0 y "Observar aves" = 1 y "Visitar Ruinas" = 1 Y  
"Visitar museos" = 0

Esto indica que el usuario, de las 4 actividades que tienen disponibles (para este ejemplo), ha seleccionado Observar Aves y Visitar Ruinas.

Aplicando el clasificador de Naive Bayes, como se muestra a continuación, se obtienen los siguientes resultados:

$$v_{\text{nb}} = \arg \max_{v_j \in V} P(v_j) \prod_i P(a_i | v_j)$$

$$P(R = T1) = P(T = 1) \times P(BM = 0 / Tipo = 1) \times P(OA = 1 / Tipo = 1) \times \\ P(VR = 1 / Tipo = 1) \times P(VM = 0 / Tipo = 1)$$

$$P(R = T1) = 0,222 \times 1 \times 0 \times 0,5 \times 1 = 0$$

$$P(R = T3) = P(T = 3) \times P(BM = 0 / Tipo = 3) \times P(OA = 1 / Tipo = 3) \times \\ P(VR = 1 / Tipo = 3) \times P(VM = 0 / Tipo = 3)$$

$$P(R = T3) = 0,444 \times 0,75 \times 0,25 \times 0,25 \times 0,5 = 0,010417$$

$$P(R = T4) = P(T = 4) \times P(BM = 0 / Tipo = 4) \times P(OA = 1 / Tipo = 4) \times \\ P(VR = 1 / Tipo = 4) \times P(VM = 0 / Tipo = 4)$$

$$P(R = T4) = 0,333 \times 0,667 \times 0,333 \times 0,667 \times 0,333 = 0,0164$$

Con los resultados de estas operaciones se obtiene el vector de probabilidad correspondiente al recorrido generado por las variables seleccionadas:

Vector de probabilidades	
Tipo	Probabilidad
1	0,0000
3	0,0104
4	0,0164

De esta vector, aplicando el clasificador bayesiano, se toma el argumento máximo, concluyendo que el tipo de atractivo requerido es de tipo 4.

Como se puede observar, cada vez que se realiza un recorrido, se calcula un vector de probabilidades, y a partir de este, se determina el tipo de atractivo que se está requiriendo.

Este método no elimina variables y permite escoger todas las actividades posibles, brindando mayor flexibilidad al usuario.



## ANEXO 10D

### IMPLEMENTACIÓN DE LA FUNCIÓN LINEAL DISCRIMINANTE

Para demostrar la aplicación de las funciones lineales discriminantes utilizadas, en el sistema de sitios turísticos, se han tomado como datos de entrada la siguiente matriz, que contiene el promedio de calificaciones de cada atractivo:

ID Atractivo	Calificación Atractivo	Calificación Seguridad	Calificación Limpieza
1	3,00	2,00	3,00
2	1,00	1,00	1,00
3	3,00	5,00	2,00
4	2,00	1,00	3,00
5	5,00	3,50	2,50
6	3,00	5,00	4,00
7	2,71	2,21	2,57
8	3,00	1,00	2,00
9	2,00	2,00	2,00
10	1,00	2,00	3,00
11	3,00	1,00	3,00
12	2,67	2,33	2,33
13	4,00	3,50	4,00
14	2,00	4,00	5,00
15	3,00	2,00	2,00
16	3,00	3,00	4,00

Se asignan los coeficientes de probabilidad para cada función discriminante lineal, aplicada a cada calificación dada al atractivo, como se muestra en las siguientes tablas:

Coeficientes de Probabilidad (Atractivo)	
$a_1$	0,7
$a_2$	0,15
$a_3$	0,15

Coeficientes de Probabilidad (Seguridad)	
$b_1$	0,15
$b_2$	0,7
$b_3$	0,15

Coeficientes de Probabilidad (Limpieza)	
$c_1$	0,15
$c_2$	0,15
$c_3$	0,70

Luego se evalúan cada una de las funciones mostradas a continuación:

$$f(\text{atractivo}) = x_1a_1 + x_2a_2 + x_3a_3$$

$$f(\text{seguridad}) = x_1b_1 + x_2b_2 + x_3b_3$$

$$f(\text{lim pieza}) = x_1c_1 + x_2c_2 + x_3c_3$$

Y se obtienen los resultados que se muestran en la siguiente tabla, que indican en que tipo de calificación encaja cada atractivo:

Atractivo	f(atractivo)	f(seguridad)	f(limpieza)	Valor Máximo	Tipo
1	2,85	2,30	2,85	2,85	Limpieza
2	1,00	1,00	1,00	1,00	Limpieza
3	3,15	4,25	2,60	4,25	Seguridad
4	2,00	1,45	2,55	2,55	Limpieza
5	4,40	3,58	3,03	4,40	Total
6	3,45	4,55	4,00	4,55	Seguridad
7	2,62	2,34	2,54	2,62	Total

8	2,55	1,45	2,00	2,55	Total
9	2,00	2,00	2,00	2,00	Limpieza
10	1,45	2,00	2,55	2,55	Limpieza
11	2,70	1,60	2,70	2,70	Limpieza
12	2,57	2,38	2,38	2,57	Total
13	3,93	3,65	3,93	3,93	Limpieza
14	2,75	3,85	4,40	4,40	Limpieza
15	2,70	2,15	2,15	2,70	Total
16	3,15	3,15	3,70	3,70	Limpieza

De los resultados obtenidos de la tabla anterior se escoge aquel que tenga el valor máximo para cada grupo, y con esto se genera el ranking de atractivos:

Id_Atractivo	f(atractivo)	f(seguridad)	f(limpieza)	Valor Máximo	Tipo
14	2,75	3,85	4,40	<b>4,40</b>	<b>Limpieza</b>
13	3,93	3,65	3,93	3,93	Limpieza
16	3,15	3,15	3,70	3,70	Limpieza
1	2,85	2,30	2,85	2,85	Limpieza
11	2,70	1,60	2,70	2,70	Limpieza
4	2,00	1,45	2,55	2,55	Limpieza
10	1,45	2,00	2,55	2,55	Limpieza
9	2,00	2,00	2,00	2,00	Limpieza
2	1,00	1,00	1,00	1,00	Limpieza
<b>6</b>	<b>3,45</b>	<b>4,55</b>	<b>4,00</b>	<b>4,55</b>	<b>Seguridad</b>
3	3,15	4,25	2,60	4,25	Seguridad
<b>5</b>	<b>4,40</b>	<b>3,58</b>	<b>3,03</b>	<b>4,40</b>	<b>Atractivo</b>
15	2,70	2,15	2,15	2,70	Atractivo
7	2,62	2,34	2,54	2,62	Atractivo
12	2,57	2,38	2,38	2,57	Atractivo
8	2,55	1,45	2,00	2,55	Atractivo

Esto indica que: el atractivo con mejor calificación general es aquel cuyo id\_Atractivo es "5"; el atractivo con mejor calificación en limpieza es el que corresponde al id\_Atractivo igual a "14"; y, el atractivo con mejor calificación en seguridad es el que posee id\_Atractivo igual a "6".

# ANEXO 11

## RESULTADOS DE LA IMPLEMENTACIÓN DEL ALGORITMO DE NAIVE BAYES

Para demostrar los resultados de la implementación del algoritmo de Naive Bayes en el sistema, se realiza el siguiente recorrido, en el que se escogen las actividades: practicar buceo, realizar parapente y visitar ruinas arqueológicas; junto con las características del sitio: lugar para comer=si, parqueo=no importa, seguridad=media, limpieza=no importa y concurrencia=alta, como se muestra a continuación:

### Búsqueda por preferencias turísticas del visitante

**Características de Lugar a Buscar**

**Lugar Para Comer:** Si

**Parqueo:** No Importa

**Seguridad:** Media

**Limpieza:** No Importa

**Concurrencia:** Alta

**Actividades**

<p style="color: #4F81BD; margin: 0;"><i>Deportes</i></p> <p><input type="checkbox"/> Jugar basket</p> <p><input type="checkbox"/> Jugar bolos</p> <p><input type="checkbox"/> Jugar fútbol</p> <p><input type="checkbox"/> Jugar volley</p> <p><input type="checkbox"/> Nadar</p> <p><input checked="" type="checkbox"/> Practicar buceo</p> <p><input type="checkbox"/> Practicar fútbol de playa</p>	<p style="color: #4F81BD; margin: 0;"><i>Deportes Extremos</i></p> <p><input type="checkbox"/> Motos acuáticas</p> <p><input type="checkbox"/> Practicar body-boarding</p> <p><input type="checkbox"/> Practicar motocross</p> <p><input type="checkbox"/> Practicar sky acuático</p> <p><input checked="" type="checkbox"/> Realizar Parapente</p> <p><input type="checkbox"/> Surfear</p> <p><input type="checkbox"/> Tablas de velas (windsurf)</p> <p><input type="checkbox"/> Volar en alas delta</p>	<p style="color: #4F81BD; margin: 0;"><i>Diversión Nocturna</i></p> <p><input type="checkbox"/> Ir a bares</p> <p><input type="checkbox"/> Ir a discotecas</p> <p><input type="checkbox"/> Ir a una peña</p>
<p style="color: #4F81BD; margin: 0;"><i>Cultura y Ciencia</i></p> <p><input type="checkbox"/> Visitar comunidades indígenas</p> <p><input type="checkbox"/> Visitar Iglesias</p> <p><input type="checkbox"/> Visitar monumentos históricos</p> <p><input type="checkbox"/> Visitar museos</p> <p><input checked="" type="checkbox"/> Visitar ruinas arqueológicas</p> <p><input type="checkbox"/> Visitar teatros</p>	<p style="color: #4F81BD; margin: 0;"><i>Ecoturismo</i></p> <p><input type="checkbox"/> Admirar fauna</p> <p><input type="checkbox"/> Admirar flora</p> <p><input type="checkbox"/> Conocer estuarios</p> <p><input type="checkbox"/> Conocer manglares</p> <p><input type="checkbox"/> Observar delfines</p> <p><input type="checkbox"/> Observar aves</p>	<p style="color: #4F81BD; margin: 0;"><i>Gastronomía</i></p> <p><input type="checkbox"/> Degustar platos típicos</p> <p><input type="checkbox"/> Visitar cafeterías</p>

En la primera fase se preseleccionan todos los atractivos en los que se puede realizar al menos una de las actividades escogidas:

<b>Paso 1: Atractivos preseleccionados</b>			
<b>Id_ Atractivo</b>	<b>Ciudad</b>	<b>Atractivo</b>	<b>Actividad</b>
2	Bahía de Caráquez	Playa de Bahía de Caráquez	Realizar Parapente
4	Bahía de Caráquez	Zona Arqueológica de Chirije	Visitar ruinas arqueológicas
9	Canoa	Playa de Canoa	Realizar Parapente
11	Crucita	Playa de Crucita	Realizar Parapente
12	Esmeraldas	Selva húmeda tropical del Chocó Biogeográfico	Visitar ruinas arqueológicas
27	Jama	Playa de Jama	Visitar ruinas arqueológicas
31	La Tolita	Isla La Tolita Pampa de Oro	Visitar ruinas arqueológicas
34	Manta	Playas de Manta	Realizar Parapente
35	Manta	Museo Centro Cultural de Manta	Visitar ruinas arqueológicas
36	Manta	Montecristi	Visitar ruinas arqueológicas
51	Puerto López	Parque Nacional de Machalilla	Visitar ruinas arqueológicas
52	Puerto López	Playa de Puerto López	Visitar ruinas arqueológicas
55	Puerto López	Isla de la Plata	Visitar ruinas arqueológicas
56	Puerto López	Museo In Situ de Agua Blanca	Visitar ruinas arqueológicas
59	Puerto López	Museo de Salango	Visitar ruinas arqueológicas
60	Rocafuerte	Rocafuerte	Visitar ruinas arqueológicas
64	San Jacinto-San Clemente	Playa San Jacinto y San Clemente	Visitar ruinas arqueológicas
87	Puerto López	Los Frailes	Practicar buceo

De estos atractivos se genera la tabla de datos a partir de la cual se realiza la preselección del tipo de atractivo:

Realizar Parapente	Visitar ruinas arqueológicas	Practicar buceo	TIPO
1	0	0	2
0	1	0	3
1	0	0	2
1	0	0	2
0	1	0	6
0	1	0	2
0	1	0	6
1	0	0	2
0	1	0	3
0	1	0	1
0	1	0	6
0	1	0	2
0	1	0	6
0	1	0	3
0	1	0	3
0	1	0	1
0	1	0	2
1	0	0	2
0	0	1	2

Una vez aplicado el algoritmo de Naive Bayes, se genera el siguiente vector de probabilidades, del cual se preselecciona el tipo de atractivo con mayor probabilidad, en este caso, el 2.

Vector de Probabilidades				
Tipo de Atractivo				
1	2	3	4	6
0,397849	0,645444	0,462366	0,021505	0,49063

Para la segunda fase, se escogen todos los atractivos que sean del tipo de atractivo preseleccionado, para aplicar el algoritmo; tomando como variables las **características del atractivo** y como la variable clase **ciudad**, obteniendo la siguiente tabla:

Lugar para comer	Parqueo	Seguridad	Limpieza	Concurrencia	Id_Ciudad
1	1	2	3	2	5
1	2	2	3	3	6
2	2	2	2	1	6
1	2	3	3	1	6
2	1	2	1	1	12
1	2	2	3	1	12
1	2	2	1	1	12
1	2	1	2	2	21
1	2	2	2	2	108
1	2	1	2	2	109
1	2	1	2	2	109
1	1	3	2	2	110
1	1	1	2	3	113
1	1	3	3	2	115
1	2	3	1	1	118
1	1	2	2	3	119
1	2	1	2	2	122
1	1	2	2	3	126
1	1	1	2	1	130
1	1	2	2	3	134
1	2	1	1	2	134
2	1	3	3	2	134
1	2	2	3	2	144
1	2	2	3	2	144
2	2	2	2	2	144
1	2	1	1	2	151
1	1	3	3	3	162
1	1	2	2	2	163
1	1	2	2	3	173
1	1	3	1	3	177
1	2	3	2	2	177
1	2	3	2	2	181
2	2	2	2	2	181
1	2	2	2	3	181

1	2	3	2	3	<b>181</b>
1	1	2	3	3	<b>181</b>
1	2	1	1	3	<b>181</b>

Al aplicar nuevamente Naive bayes se obtiene el siguiente vector de probabilidades, del que se escoge la primera ciudad cuyo valor de probabilidad sea mayor, que para este caso es la ciudad 119 que corresponde a Crucita, como se muestra en la siguiente tabla.


Ciudad	Id_ciudad	Probabilidad
Atacames	5	0,203
Muisne	6	0,175
Tonchigue	12	0,175
San Jacinto-San Clemente	21	0,103
Same	108	0,203
San Lorenzo	109	0,105
Sua	110	0,103
Bahía de Caráquez	113	0,203
Canoa	115	0,103
Cojimíes	118	0,103
<b>Crucita</b>	<b>119</b>	<b>0,303</b>
Jama	122	0,103
Manta	126	0,303
Pedernales	130	0,103
Puerto López	134	0,141
Babahoyo	144	0,175
Vinces	151	0,103
Guayaquil	162	0,203
La Libertad	163	0,203
General Villamil	173	0,303
Salinas	177	0,155
Santa Elena	181	0,216




Adicionalmente, el sistema sugiere las tres ciudades siguientes con mayor probabilidad en orden descendente, como se aprecia en la siguiente ventana, que muestra los resultados de la búsqueda por perfil.

### Resultado de Búsqueda

La ciudad que mejor se asemeja a sus preferencias es:

<b>Puesto</b> <b>#</b> <b>1</b>	<b>Descripción:</b> <u>Crucita</u> Denominada «Playa, Sol y Mar del Ecuador», por su incomparable belleza, Crucita constituye un escenario natural propio para la realización de importantes y exitosos eventos de belleza, que constantemente atraen hasta sus playas a miles y miles de turistas <b>Provincia</b> <u>Manabí</u>	
---------------------------------------	---	---

Los atractivos que puede encontrar en Crucita son:

Atractivos	Descripción	Imagen
<u>Playa de Crucita</u>	Es una de las más bellas playas del Ecuador. Por eso se la conoce también con el nombre de Playa "La Bella". Es el destino más visitado por los fanáticos del deporte de aventura que desean practicar el parapente y alas delta. Crucita es considerado uno de los mejores lugares para realizar este deporte en Ecuador.	

Otras Opciones para sus preferencias son:

Puesto	Ciudad
2	<u>Manta</u>
3	<u>General Villamil</u>
4	<u>Santa Elena</u>

## ANEXO 12

### VECTORES DE PROBABILIDAD DE CADA ACTIVIDAD

Bañarse en el mar						
Valor	Tipo Atractivo					
	1	2	3	4	5	6
1	0	0,333333	0,083333	0	0	0
0	1	0,666667	0,916667	1	1	1

Montar a caballo						
Valor	Tipo Atractivo					
	1	2	3	4	5	6
0	1	0,969697	1	1	1	1
1	0	0,030303	0	0	0	0

Navegar en botes						
Valor	Tipo Atractivo					
	1	2	3	4	5	6
1	0	0,151515	0	0	0	0,060606
0	1	0,848485	1	1	1	0,939394

Acampar						
Valor	Tipo Atractivo					
	1	2	3	4	5	6
0	1	1	1	1	1	1

Excursionar						
Valor	Tipo Atractivo					
	1	2	3	4	5	6
0	1	1	1	1	1	0,939394
1	0	0	0	0	0	0,060606

Ir a bares						
Valor	Tipo Atractivo					
	1	2	3	4	5	6
0	1	0,878788	1	1	1	1
1	0	0,121212	0	0	0	0

Ir a discotecas						
Valor	Tipo Atractivo					
	1	2	3	4	5	6
0	1	0,878788	1	1	1	1

1	0	0,121212	0	0	0	0
---	---	----------	---	---	---	---

Ir de compras						
Valor	Tipo Atractivo					
	1	2	3	4	5	6
0	1	1	1	1	1	1

Jugar basket						
Valor	Tipo Atractivo					
	1	2	3	4	5	6
0	1	1	1	1	1	1

Observar aves						
Valor	Tipo Atractivo					
	1	2	3	4	5	6
0	1	0,878788	0,916667	1	1	0,575758
1	0	0,121212	0,083333	0	0	0,424242

Tomar el sol						
Valor	Tipo Atractivo					
	1	2	3	4	5	6
1	0	0,30303	0,083333	0	0	0,030303
0	1	0,69697	0,916667	1	1	0,969697

Jugar volley						
Valor	Tipo Atractivo					
	1	2	3	4	5	6
1	0	0,151515	0	0	0	0
0	1	0,848485	1	1	1	1

Surfear						
Valor	Tipo Atractivo					
	1	2	3	4	5	6
1	0	0,212121	0	0	0	0,030303
0	1	0,787879	1	1	1	0,969697

Practicar motocross						
Valor	Tipo Atractivo					
	1	2	3	4	5	6
0	1	1	1	1	1	1

Nadar						
Valor	Tipo Atractivo					
	1	2	3	4	5	6
1	0	0,212121	0,083333	0	0	0,030303
0	1	0,787879	0,916667	1	1	0,969697

Realizar Parapente						
Valor	Tipo Atractivo					
	1	2	3	4	5	6
0	1	0,939394	1	1	1	1
1	0	0,060606	0	0	0	0

Ir a una peña						
Valor	Tipo Atractivo					
	1	2	3	4	5	6
1	0	0,060606	0	0	0	0
0	1	0,939394	1	1	1	1

Ir de pesca						
Valor	Tipo Atractivo					
	1	2	3	4	5	6
1	0	0,151515	0	0	0	0,060606
0	1	0,848485	1	1	1	0,939394

Practicar sky acuático						
Valor	Tipo Atractivo					
	1	2	3	4	5	6
0	1	1	1	1	1	0,969697
1	0	0	0	0	0	0,030303

Pasear en bananas						
Valor	Tipo Atractivo					
	1	2	3	4	5	6
1	0	0,030303	0	0	0	0
0	1	0,969697	1	1	1	1

Alquilar botes						
Valor	Tipo Atractivo					
	1	2	3	4	5	6
0	1	0,909091	1	1	1	0,969697
1	0	0,090909	0	0	0	0,030303

Tablas de velas (windsurf)						
Valor	Tipo Atractivo					
	1	2	3	4	5	6
0	1	0,969697	1	1	1	1
1	0	0,030303	0	0	0	0

Observar ballenas jorobadas						
Valor	Tipo Atractivo					
	1	2	3	4	5	6

<b>0</b>	1	0,939394	1	1	1	0,939394
<b>1</b>	0	0,060606	0	0	0	0,060606

<b>Observar delfines</b>						
<b>Valor</b>	<b>Tipo Atractivo</b>					
	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>
<b>0</b>	1	0,969697	1	1	1	0,969697
<b>1</b>	0	0,030303	0	0	0	0,030303

<b>Andar en triciclos</b>						
<b>Valor</b>	<b>Tipo Atractivo</b>					
	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>
<b>0</b>	1	0,939394	1	1	1	1
<b>1</b>	0	0,060606	0	0	0	0

<b>Motos acuáticas</b>						
<b>Valor</b>	<b>Tipo Atractivo</b>					
	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>
<b>0</b>	1	0,969697	1	1	1	1
<b>1</b>	0	0,030303	0	0	0	0

<b>Observar vuelo de gaviotas</b>						
<b>Valor</b>	<b>Tipo Atractivo</b>					
	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>
<b>0</b>	1	1	1	1	1	0,878788
<b>1</b>	0	0	0	0	0	0,121212

<b>Observar crustáceos</b>						
<b>Valor</b>	<b>Tipo Atractivo</b>					
	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>
<b>0</b>	1	0,969697	1	1	1	0,909091
<b>1</b>	0	0,030303	0	0	0	0,090909

<b>Observar conchas</b>						
<b>Valor</b>	<b>Tipo Atractivo</b>					
	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>
<b>0</b>	1	1	1	1	1	0,969697
<b>1</b>	0	0	0	0	0	0,030303

<b>Observar garzas bueyeras</b>						
<b>Valor</b>	<b>Tipo Atractivo</b>					
	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>
<b>0</b>	1	1	1	1	1	1

<b>Observar piqueros de patas azules</b>						
<b>Valor</b>	<b>Tipo Atractivo</b>					

	1	2	3	4	5	6
0	1	0,969697	1	1	1	0,969697
1	0	0,030303	0	0	0	0,030303

Observar Peñascos						
Valor	Tipo Atractivo					
	1	2	3	4	5	6
0	1	0,878788	1	1	1	0,969697
1	0	0,121212	0	0	0	0,030303

Visitar ruinas arqueológicas						
Valor	Tipo Atractivo					
	1	2	3	4	5	6
0	1	1	0,916667	1	1	0,939394
1	0	0	0,083333	0	0	0,060606

Conocer manglares						
Valor	Tipo Atractivo					
	1	2	3	4	5	6
0	1	0,939394	1	1	1	0,787879
1	0	0,060606	0	0	0	0,212121

Visitar haciendas						
Valor	Tipo Atractivo					
	1	2	3	4	5	6
0	1	1	1	1	1	1

Visitar reservas de bosques						
Valor	Tipo Atractivo					
	1	2	3	4	5	6
0	1	0,969697	0,916667	1	1	0,606061
1	0	0,030303	0,083333	0	0	0,393939

Practicar body-boarding						
Valor	Tipo Atractivo					
	1	2	3	4	5	6
0	1	1	1	1	1	1

Degustar platos típicos						
Valor	Tipo Atractivo					
	1	2	3	4	5	6
0	1	0,939394	0,916667	1	1	1
1	0	0,060606	0,083333	0	0	0

Volar en alas delta						
Valor	Tipo Atractivo					

	1	2	3	4	5	6
0	1	0,909091	1	1	1	1
1	0	0,090909	0	0	0	0

Descansar en hamacas						
Valor	Tipo Atractivo					
	1	2	3	4	5	6
1	0	0,060606	0	0	0	0
0	1	0,939394	1	1	1	1

Visitar Iglesias						
Valor	Tipo Atractivo					
	1	2	3	4	5	6
1	0	0,030303	0	0	0	0
0	1	0,969697	1	1	1	1

Conocer estuarios						
Valor	Tipo Atractivo					
	1	2	3	4	5	6
0	1	0,939394	1	1	1	0,939394
1	0	0,060606	0	0	0	0,060606

Visitar museos						
Valor	Tipo Atractivo					
	1	2	3	4	5	6
0	1	1	0,833333	1	1	0,969697
1	0	0	0,166667	0	0	0,030303

Pasear por el malecón						
Valor	Tipo Atractivo					
	1	2	3	4	5	6
1	0	0,060606	0	0	0	0
0	1	0,939394	1	1	1	1

Alquilar cabañas						
Valor	Tipo Atractivo					
	1	2	3	4	5	6
0	1	0,939394	0,916667	1	1	1
1	0	0,060606	0,083333	0	0	0

Admirar fauna						
Valor	Tipo Atractivo					
	1	2	3	4	5	6
0	1	0,848485	0,916667	1	1	0,545455
1	0	0,151515	0,083333	0	0	0,454545

Visitar parques nacionales						
Valor	Tipo Atractivo					
	1	2	3	4	5	6
0	1	1	1	1	1	1

Observar orquídeas						
Valor	Tipo Atractivo					
	1	2	3	4	5	6
0	1	1	1	1	1	1

Recolectar crustáceos						
Valor	Tipo Atractivo					
	1	2	3	4	5	6
0	1	1	1	1	1	1

Practicar senderismo						
Valor	Tipo Atractivo					
	1	2	3	4	5	6
0	1	1	0,916667	1	1	0,939394
1	0	0	0,083333	0	0	0,060606

Visitar ríos						
Valor	Tipo Atractivo					
	1	2	3	4	5	6
0	1	1	1	1	1	0,939394
1	0	0	0	0	0	0,060606

Visitar cerros						
Valor	Tipo Atractivo					
	1	2	3	4	5	6
0	1	1	0,916667	1	1	0,939394
1	0	0	0,083333	0	0	0,060606

Visitar monumentos históricos						
Valor	Tipo Atractivo					
	1	2	3	4	5	6
0	1	1	1	1	1	1

Visitar esteros						
Valor	Tipo Atractivo					
	1	2	3	4	5	6
0	1	1	1	1	1	0,939394
1	0	0	0	0	0	0,060606

Comprar artesanías						
Valor	Tipo Atractivo					



	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>
<b>1</b>	0	0,060606	0	0	0	0
<b>0</b>	1	0,939394	1	1	1	1

<b>Cantar karaoke</b>						
<b>Valor</b>	<b>Tipo Atractivo</b>					
	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>
<b>0</b>	1	1	1	1	1	1

<b>Ingerir bebidas alcohólicas</b>						
<b>Valor</b>	<b>Tipo Atractivo</b>					
	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>
<b>0</b>	1	1	1	1	1	1

<b>Pasear a caballo en la playa</b>						
<b>Valor</b>	<b>Tipo Atractivo</b>					
	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>
<b>0</b>	1	0,969697	1	1	1	1
<b>1</b>	0	0,030303	0	0	0	0

<b>Escuchar música en vivo</b>						
<b>Valor</b>	<b>Tipo Atractivo</b>					
	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>
<b>0</b>	1	1	1	1	1	1

<b>Visitar teatros</b>						
<b>Valor</b>	<b>Tipo Atractivo</b>					
	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>
<b>0</b>	1	1	1	1	1	1

<b>Visitar parques infantiles</b>						
<b>Valor</b>	<b>Tipo Atractivo</b>					
	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>
<b>0</b>	1	1	1	1	1	1

<b>Visitar casinos y bingos</b>						
<b>Valor</b>	<b>Tipo Atractivo</b>					
	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>
<b>0</b>	1	1	1	1	1	1

<b>Jugar fútbol</b>						
<b>Valor</b>	<b>Tipo Atractivo</b>					
	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>
<b>0</b>	1	1	1	1	1	1

<b>Practicar fútbol de playa</b>						
----------------------------------	--	--	--	--	--	--

Valor	Tipo Atractivo					
	1	2	3	4	5	6
0	1	0,969697	1	1	1	1
1	0	0,030303	0	0	0	0

Observar tortugas						
Valor	Tipo Atractivo					
	1	2	3	4	5	6
0	1	0,969697	1	1	1	0,939394
1	0	0,030303	0	0	0	0,060606

Visitar comunidades indígenas						
Valor	Tipo Atractivo					
	1	2	3	4	5	6
0	1	1	1	1	1	0,969697
1	0	0	0	0	0	0,030303

Admirar flora						
Valor	Tipo Atractivo					
	1	2	3	4	5	6
0	1	0,969697	1	1	1	0,636364
1	0	0,030303	0	0	0	0,363636

Visitar cafeterías						
Valor	Tipo Atractivo					
	1	2	3	4	5	6
0	1	1	1	1	1	1

Jugar bolos						
Valor	Tipo Atractivo					
	1	2	3	4	5	6
0	1	1	1	1	1	1

Tipo Atractivo	Probabilidad
1	0,058140
2	0,383721
3	0,139535
4	0,023256
5	0,011628
6	0,383721

## **ANEXO 13**

### **MATRIZ DE VARIABLES USADAS PARA LA PREDICCIÓN**



