



Esta obra está bajo una [Licencia Creative Commons Atribución- NoComercial-Compartirigual 2.5 Perú](http://creativecommons.org/licenses/by-nc-sa/2.5/pe/).  
Vea una copia de esta licencia en <http://creativecommons.org/licenses/by-nc-sa/2.5/pe/>



**UNIVERSIDAD NACIONAL DE SAN MARTÍN – TARAPOTO**  
**FACULTAD DE INGENIERÍA DE SISTEMAS E INFORMÁTICA**  
**ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS E**  
**INFORMÁTICA**



**Sistema basado en el modelo de recuperación vectorial para el servicio de  
recomendación de restaurantes en la ciudad de Tarapoto**

**Tesis para optar por el título profesional de Ingeniero de Sistemas e  
Informática**

**AUTOR:**

**Bach. Alicia Lizbeth Rafael Saldaña**

**ASESOR:**

**Ing. M. Sc. Pedro Antonio Gonzales Sánchez**

**Tarapoto - Perú**

**2018**

**UNIVERSIDAD NACIONAL DE SAN MARTÍN – TARAPOTO**  
**FACULTAD DE INGENIERÍA DE SISTEMAS E INFORMÁTICA**  
**ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS E**  
**INFORMÁTICA**

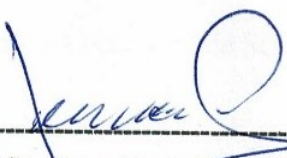



**Sistema basado en el modelo de recuperación vectorial para el servicio de  
recomendación de restaurantes en la ciudad de Tarapoto**

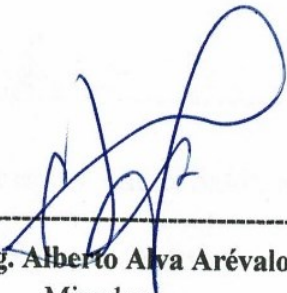
**AUTOR:**

**Bach. Alicia Lizbeth Rafael Saldaña**

**Sustentada y aprobada ante el honorable jurado el día 17 de octubre del 2018**

  
-----  
**Ing. Mg. Juan Carlos García Castro**  
Presidente

  
-----  
**Ing. Carlos Francois Hidalgo Reátegui**  
Secretario

  
-----  
**Ing. Mg. Alberto Alva Arévalo**  
Miembro

## Declaratoria de Autenticidad

Yo, Alicia Lizbeth Rafael Saldaña, bachiller de la Facultad de Ingeniería de Sistemas e Informática, de la Escuela profesional de Ingeniería de Sistemas e Informática de la Universidad Nacional de San Martín – Tarapoto, identificado con DNI N° 70652042, con la tesis titulada “Sistema basado en el modelo de recuperación vectorial para el servicio de recomendación de restaurantes en la ciudad de Tarapoto”.

Declaro bajo juramento que:

1. La tesis presentada es de mi autoría.
2. He respetado las normas internacionales de citas y referencias para las fuentes consultadas. Por tanto, la tesis no ha sido plagiada ni total ni parcialmente.
3. La tesis no ha sido auto plagiada; es decir, no ha sido publicada ni presentada anteriormente para obtener algún grado académico previo o título profesional.
4. Los datos presentados en los resultados son reales, no han sido falseados, ni duplicados, ni copiados y por tanto los resultados que se presenten en la tesis se constituirán en aportes a la realidad investigada.

De considerar que el trabajo cuenta con una falta grave, como el hecho de contar con datos fraudulentos, demostrar indicios y plagio (al no citar la información con sus autores). Plagio (al presentar información de otros trabajos como propios), falsificación (al presentar la información de ideas de otras personas de forma falsa), entre otros, asumo las consecuencias y sanciones que de mi acción se deriven, sometiéndome a la normatividad vigente de la Universidad Nacional de San Martín – Tarapoto.

Tarapoto, 17 de Octubre del 2018



Alicia Lizbeth Rafael Saldaña

DNI N°70652042

**Formato de autorización NO EXCLUSIVA para la publicación de trabajos de investigación, conducentes a optar grados académicos y títulos profesionales en el Repositorio Digital de Tesis.**

**1. Datos del autor:**

Apellidos y nombres: <i>Rafael Saldana, Alicia Lizbeth</i>	
Código de alumno : <i>127146</i>	Teléfono: <i>959607618</i>
Correo electrónico : <i>alirs.ing@gmail.com</i>	DNI: <i>70652042</i>

(En caso haya más autores, llenar un formulario por autor)

**2. Datos Académicos**

Facultad de: <i>Ingeniería de Sistemas e Informática</i>
Escuela Profesional de: <i>Ingeniería de Sistemas e Informática</i>

**3. Tipo de trabajo de investigación**

Tesis	(x)	Trabajo de investigación	( )
Trabajo de suficiencia profesional	( )		

**4. Datos del Trabajo de investigación**

Título: <i>Sistema basado en el modelo de recuperación vectorial para el servicio de recomendación de restaurantes en la ciudad de Tarapoto</i>
Año de publicación: <i>2018</i>

**5. Tipo de Acceso al documento**

Acceso público *	(x)	Embargo	( )
Acceso restringido **	( )		

Si el autor elige el tipo de acceso abierto o público, otorga a la Universidad Nacional de San Martín – Tarapoto, una licencia **No Exclusiva**, para publicar, conservar y sin modificar su contenido, pueda convertirla a cualquier formato de fichero, medio o soporte, siempre con fines de seguridad, preservación y difusión en el Repositorio de Tesis Digital. Respetando siempre los Derechos de Autor y Propiedad Intelectual de acuerdo y en el Marco de la Ley 822.

En caso que el autor elija la segunda opción, es necesario y obligatorio que indique el sustento correspondiente:


**6. Originalidad del archivo digital.**

Por el presente dejo constancia que el archivo digital que entrego a la Universidad Nacional de San Martín - Tarapoto, como parte del proceso conducente a obtener el título profesional o grado académico, es la versión final del trabajo de investigación sustentado y aprobado por el Jurado.

## 7. Otorgamiento de una licencia **CREATIVE COMMONS**

Para investigaciones que son de acceso abierto se les otorgó una licencia *Creative Commons*, con la finalidad de que cualquier usuario pueda acceder a la obra, bajo los términos que dicha licencia implica

<https://creativecommons.org/licenses/by-nc-sa/2.5/pe/>

El autor, por medio de este documento, autoriza a la Universidad Nacional de San Martín - Tarapoto, publicar su trabajo de investigación en formato digital en el Repositorio Digital de Tesis, al cual se podrá acceder, preservar y difundir de forma libre y gratuita, de manera íntegra a todo el documento.

Según el inciso 12.2, del artículo 12° del Reglamento del Registro Nacional de Trabajos de Investigación para optar grados académicos y títulos profesionales - RENATI “**Las universidades, instituciones y escuelas de educación superior tienen como obligación registrar todos los trabajos de investigación y proyectos, incluyendo los metadatos en sus repositorios institucionales precisando si son de acceso abierto o restringido, los cuales serán posteriormente recolectados por el Repositorio Digital RENATI, a través del Repositorio ALICIA**”.



.....  
Firma del Autor

---

## 8. Para ser llenado en la Oficina de Repositorio Digital de Ciencia y Tecnología de Acceso Abierto de la UNSM – T.

Fecha de recepción del documento:

13 / 02 / 2019



.....  
Firma del Responsable de Repositorio  
Digital de Ciencia y Tecnología de Acceso  
Abierto de la UNSM – T.

\***Acceso abierto:** uso lícito que confiere un titular de derechos de propiedad intelectual a cualquier persona, para que pueda acceder de manera inmediata y gratuita a una obra, datos procesados o estadísticas de monitoreo, sin necesidad de registro, suscripción, ni pago, estando autorizada a leerla, descargarla, reproducirla, distribuirla, imprimirla, buscarla y enlazar textos completos (Reglamento de la Ley No 30035).

\*\* **Acceso restringido:** el documento no se visualizará en el Repositorio.

## **Dedicatoria**

Este trabajo está dedicado:

A Dios por permitirme tener vida y salud para poder cumplir uno más de mis propósitos en mi vida.

A mis padres: Paulino Rafael Altamirano y Elva de Jesús Saldaña Tello y hermanas por el apoyo que me brindaron a lo largo de mi vida sus palabras de aliento para seguir adelante y cumpla mis metas.

A mi asesor por su tiempo, dedicación, paciencia e impulsar el desarrollo en mi formación como profesional.

A todos mis docentes por inculcarme enseñanzas para crecer a lo largo de mi vida profesional.

A mis amigos y compañeros, presentes y pasados, por compartir alegrías y tristezas

MIL GRACIAS A TODOS.

## Índice

<b>Dedicatoria .....</b>	<b>iii</b>
<b>Índice .....</b>	<b>vii</b>
<b>Índice de Figuras .....</b>	<b>ix</b>
<b>Índice de tablas .....</b>	<b>x</b>
<b>Lista de siglas y abreviaturas .....</b>	<b>xi</b>
<b>Resumen .....</b>	<b>xii</b>
<b>Abstract .....</b>	<b>xiii</b>
<b>INTRODUCCIÓN .....</b>	<b>1</b>
<b>CAPÍTULO I .....</b>	<b>2</b>
<b>REVISIÓN BIBLIOGRÁFICA .....</b>	<b>2</b>
1.1. Modelos para la recuperación de información .....	2
1.2. Modelo de recuperación vectorial .....	3
1.3. Medida de la precisión media (MAP) .....	4
1.4. Difusión de información .....	5
1.5. Restaurant .....	8
1.6. Definición de términos básicos .....	8
<b>CAPÍTULO II .....</b>	<b>10</b>
<b>MATERIALES Y MÉTODOS .....</b>	<b>10</b>
2.1. Objetivos .....	10
2.2. Sistema de hipótesis .....	10
2.3. Tipo de investigación .....	12
2.4. Nivel de investigación .....	12
2.5. Diseño de investigación .....	12
2.6. Población y muestra .....	12
2.7. Técnicas e instrumentos de recolección de datos .....	13
2.8. Técnicas de procedimiento y análisis de datos .....	14



<b>CAPITULO III .....</b>	<b>15</b>
<b>RESULTADOS Y DISCUSIÓN .....</b>	<b>15</b>
3.1. Presentación de resultados .....	15
3.2. Contrastación de hipótesis .....	18
<b>CONCLUSIONES .....</b>	<b>22</b>
<b>RECOMENDACIONES .....</b>	<b>23</b>
<b>REFERENCIAS BIBLIOGRÁFICAS .....</b>	<b>24</b>
<b>ANEXOS .....</b>	<b>26</b>

## Índice de Figuras

Figura 1: Fórmula de similitud.....	2
Figura 2: Cálculo del TF de un término.....	3
Figura 3: Fórmula del factor IDF.....	4
Figura 4: Fórmula MAP.....	4
Figura 5: Difusión documental.....	6
Figura 6: Nivel de difusión de información del Servicio de recomendación de restaurantes (Pre Test).....	16
Figura 7: Nivel de difusión de información del Servicio de recomendación de restaurantes (Pos Test).....	18
Figura 8: Región crítica .....	20
Figura 9: Gráfica de dispersión.....	21

## Índice de tablas

Tabla 1: Escala de medición de las variables.....	11
Tabla 2: Diseño pre-experimental.....	12
Tabla 3: Técnicas e instrumentos utilizados.....	13
Tabla 4: Resumen de los indicadores de la difusión de información, antes del Buscador Informático.....	15
Tabla 5: Resumen de los indicadores de la difusión de información, después del Buscador Informático.....	16
Tabla 6: Valores promediados del Pos Test.....	21

## Lista de siglas y abreviaturas

<b>UNSM-T:</b>	Universidad Nacional de San Martín - Tarapoto.
<b>FISI:</b>	Facultad de Ingeniería de Sistemas e Informática.
<b>SI:</b>	Sistema Informático

## Resumen

El presente proyecto de investigación titulado “Sistema basado en el modelo de recuperación vectorial para el servicio de recomendación de restaurantes en la ciudad de Tarapoto”, desarrollado con financiamiento de la UNSM-T el año 2017, con el objetivo de determinar la mejora del servicio de recomendación de restaurantes a través de un sistema basado en el modelo de recuperación vectorial en la ciudad de Tarapoto.

Es importante mencionar que la investigación es de tipo Aplicada – Transversal porque con los resultados obtenidos se aplica para la resolución de la problemática planteada el cual genera conocimiento nuevo, de igual manera el nivel de investigación es experimental ya que se manipula a la variable independiente con el fin de conseguir una mejora.

Además se diseñó un pre y pos test, para una muestra de 138 turistas que visitan la ciudad de Tarapoto, con la cuales permitió conocer la situación actual de un antes y un después de que se presentará el sistema de recomendación de restaurantes, utilizando el Modelo de recuperación vectorial para la construcción del Sistema Informático ([www.sanmartincome.com](http://www.sanmartincome.com)).

Palabras clave: Recuperación vectorial, turista, restaurante, buscador informático, recomendar, difusión.

## Abstract

The following research project titled as "System based on the vector recovery model for the restaurant recommendation service in the city of Tarapoto", developed with funding from the UNSM-T in 2017, with the objective of determining the improvement of the service recommendation of restaurants through a system based on the model of vector recovery in the city of Tarapoto.

It is important mentioning that this is an Applied - Transversal research type because the results were applied for the resolution of the problematic raised which generates new knowledge, in the same way the research level is experimental since the independent variable is manipulated with in order to get an improvement.

In addition, a pre and posttest was designed for a sample of 138 tourists visiting the city of Tarapoto, with which they were able to know the current situation of a before and after the restaurant recommendation system will be presented, using the Model of vector recovery for the construction of the Computer System. ([www.sanmartincome.com](http://www.sanmartincome.com)).

Keywords: Vector recovery, tourist, restaurant, computer search engine, recommend, diffusion.



## INTRODUCCIÓN

El Proyecto de Investigación titulado “Sistema basado en el modelo de recuperación vectorial para el servicio de recomendación de restaurantes en la ciudad de Tarapoto”, presentado por Bach. Alicia Lizbeth Rafael Saldaña, en el marco del Reglamento de concurso de Proyectos de Investigación Científica-Tecnológica e Innovación a Nivel de Pregrado, financiados por la UNSM-T período 2017, tuvo como propósito de determinar la mejora del servicio de recomendación de restaurantes a través de un sistema basado en el modelo de recuperación vectorial en la ciudad de Tarapoto, se desarrolló entre los meses de Agosto a Diciembre del año 2017.

Se inició recopilando información referente a los restaurantes y los turistas, luego implementamos un modelo de recuperación vectorial con la información recolectada de los diferentes restaurantes de la ciudad de Tarapoto, finalmente analizamos la influencia del sistema informático en la mejora para la recomendación de los restaurantes en la ciudad. El Estudio fue de tipo Aplicada-Transversal, a Nivel Experimental, se realizó una tanto un pre test como un pos test para la muestra de 138 turistas.

En la ciudad de Tarapoto arriban un promedio de 33761 turistas mensualmente, en lo cual el 89% desconocen de los restaurantes que existen en la ciudad ya que es la primera vez que vienen o están desinformados ya que no buscaron previamente referente a este tema.

Por tal motivo este proyecto de investigación se orientó a crear una herramienta tecnológica que ayude a mejorar este servicio, poniendo énfasis en investigar cómo implementaríamos un sistema informático basado en el modelo de recomendación vectorial que se ajuste para el fin de la recomendación de los restaurantes existentes de la ciudad de Tarapoto.

Los resultados que obtenidos en nuestro proyecto demuestran que se ha logrado mejorar el servicio de recomendación de restaurantes en la ciudad de Tarapoto.

El presente informe final de este proyecto de investigación, consta de tres (3) capítulos, el primero referido las Referencias Bibliográficas, el segundo a los Materiales y Métodos, y el tercero a los Resultados de la Investigación.

Finalmente, con la investigación se espera, haber contribuido con nuevo conocimiento, beneficiando directa e indirectamente al turismo de la ciudad de Tarapoto.

# CAPÍTULO I

## REVISIÓN BIBLIOGRÁFICA

### 1.1. Modelos para la recuperación de información

#### 1.1.1. Modelo Booleano

Es un modelo de recuperación simple, basado en la teoría de conjuntos y el álgebra booleana. Dada su inherente simplicidad y su pulcro formalismo ha recibido gran atención y sido adoptado por muchos de los primeros sistemas bibliográficos comerciales. Su estrategia de recuperación está basada en un criterio de decisión binario (pertinente o no pertinente) sin ninguna noción de escala de medida, sin noción de un emparejamiento parcial en las condiciones de la pregunta. (Martínez Méndez y *col.* 2004).

#### 1.1.2. Modelo Vectorial

Fue propuesto en el año de 1968 y continúa siendo muy utilizado en la actualidad por su eficiencia en los buenos resultados obtenidos. En este modelo las páginas web (documentos) y las consultas son vistos como vectores en un espacio vectorial n-dimensional, donde la distancia vectorial es usada como medida de similitud (Sim). Cada página web (documento) es representado como un vector de términos y cada término posee un valor asociado que indica el grado de importancia (peso) de este en una determinada página web (documento). La fórmula de similitud (Sim) se basa en un conjunto de palabras y páginas web (documentos) y está dada por:

$$\text{sim}(d, q) = \cos\theta = \frac{\sum_{i=1}^t w(i, d) \times w(i, q)}{\sqrt{\sum (w(i, d))^2} \times \sqrt{\sum (w(i, q))^2}}$$

Norma de d Norma de q

Figura 1: Formula de similitud. (Fuente: Robertson, S. 2004)

#### 1.1.3. Modelo Probabilístico

El modelo de recuperación probabilístico se basa en la equiparación probabilística, dados un documento y una pregunta, es posible calcular la probabilidad de que ese documento sea relevante para esa pregunta (Martínez Méndez y *col.* 2004)



## 1.2. Modelo de recuperación vectorial

El modelo de recuperación vectorial se basa en el grado de similitud de una consulta dada por el usuario con respecto a los documentos de la colección cuyos términos fueron ponderados mediante TF-IDF. El modelo de recuperación vectorial fue presentado por Salton en 1975 y posteriormente asentado en 1983 junto con Mc Gill y se basa en tres principios esenciales.

- La equiparación parcial, esto es, la capacidad del sistema para ordenar los resultados de una búsqueda, basado en el grado de similitud entre cada documento de la colección y la consulta.
- La ponderación de los términos en los documentos, no limitándose a señalar la presencia o ausencia de los mismos, sino adscribiendo a cada término en cada documento un número real que refleje su importancia en el documento.
- La ponderación de los términos en la consulta, de manera que el usuario puede asignar pesos a los términos de la consulta que reflejen la importancia de los mismos en relación a su necesidad informativa.

Factor TF: Term Frequency : Frecuencia de Aparición de un Término. El factor TF es la suma de todas las ocurrencias o el número de veces que aparece un término en un documento. A este tipo de frecuencia de aparición también se la denomina “Frecuencia de aparición relativa” por que atañe a un documento en concreto y no a toda la colección.

Su cálculo se efectúa una vez el texto del documento ha sido normalizado, según los procesos de depuración mencionados. Posteriormente se lleva a cabo el conteo de las veces que el término aparece presente en el documento. De hecho, se deja bien claro que es necesario calcular el TF de cada término en cada documento.

$$tf(n) = \sum_{D1} (n)$$

La frecuencia de aparición de un término (n) en un documento (D1) es la suma de las ocurrencias de dicho término

Figura 2: Cálculo del TF de un término. (Fuente: Robertson, S. 2004)

Factor IDF: Inverse Document Frequency: Frecuencia Inversa del Documento para un término

El factor IDF de un término es inversamente proporcional al número de documentos en los que aparece dicho término. Esto significa que cuanto menor sea la cantidad de documentos, así como la frecuencia absoluta de aparición del término, mayor será su factor IDF y a la inversa, cuanto mayor sea la frecuencia absoluta relativa a una alta presencia en todos los documentos de la colección, menor será su factor discriminatorio. El Factor IDF es único para cada término de la colección. El IDF de un término dado (n) se realiza aplicando el logaritmo en base 10 de N (Número total de documentos de la colección) dividido entre la “Frecuencia de documentos para un término (n) en la colección” (o lo que es lo mismo el número de documentos de la colección en los que aparece en termino (n) dado). Al valor resultante se le suma 1 para corregir los valores para los términos con IDF muy bajos (Aunque esta variación depende del sistema de recuperación).

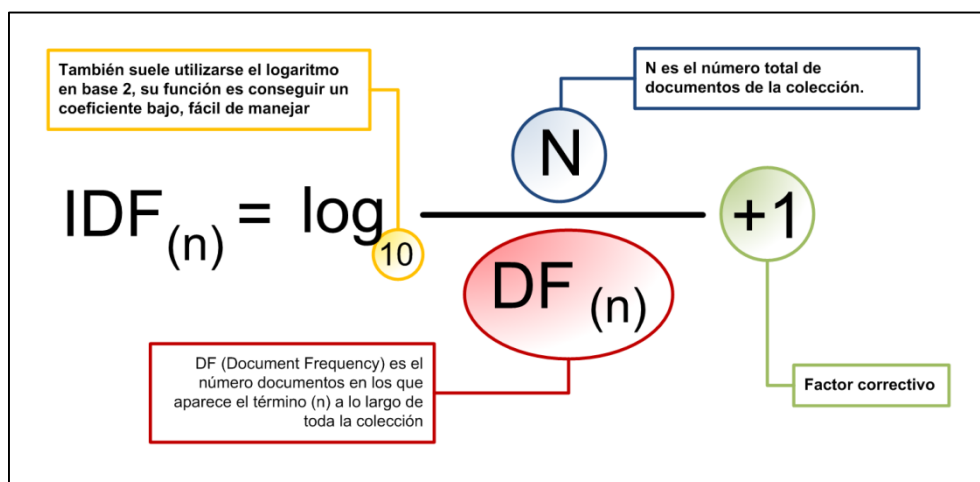


Figura 3: Fórmula del factor IDF (Fuente: Robertson, S. 2004)

### 1.3. Medida de la precisión media (MAP)

Es una medida que intenta sumar todos los valores de los promedios de las precisiones, o sea que tan preciso es el modelo de recuperación de información web para retornar páginas web (documentos) que sean relevantes o importantes para el usuario. Su fórmula está dada por (Souza, R. R. 2006):

$$MAP = \frac{\sum_{n=1}^N P@n \times rel(n)}{r_q}$$

Figura 4: Fórmula MAP. (Fuente: Souza, R. R. 2006)

## **1.4. Difusión de información**

Todo centro de documentación recoge y procesa información con el propósito de difundirla y darla a conocer. La difusión de información es el proceso por el cual se transmite al usuario la información que necesita o en darle la posibilidad de obtenerla. Se trata de una operación documental de salida. En su sentido más amplio engloba todas las modalidades de transmisión de documentos o referencias informativas. (Castillo, L. (2004-2005)).

### **1.4.1. Formas de difusión**

No hay una forma única de difusión, sino diferentes tipos de productos y servicios capaces de vehicular la información hacia los usuarios. Con objeto de sistematizar se pueden distinguir dos formas básicas de difusión, la difusión bajo demanda y la difusión documental.

En cuanto a las vías de difusión puede ser: papel impreso, tablones de anuncios, expositores, soporte magnético para consultar en ordenadores, página web, correo electrónico, difusión verbal (persona a persona, conferencias, cursos), medios audiovisuales, videos informativos. (Castillo, L. (2004-2005)).

### **1.4.2. La difusión bajo demanda**

Se trata de una difusión pasiva. La iniciativa parte del usuario que necesita información. Se realiza cuando el usuario se dirige al centro con el objetivo de solicitar una información concreta. El documentalista interroga al sistema y recupera la información solicitada. Las búsquedas bibliográficas y documentales responden a esta necesidad de información planteada por el usuario. Pueden realizarse sobre base de datos documentales creadas por el propio centro pero, también, sobre bases externas. Se considera también difusión bajo demanda la consulta en línea a una base de datos por parte del usuario sin la mediación de un documentalista. (Castillo, L. (2004-2005)).

### 1.4.3. La difusión documental

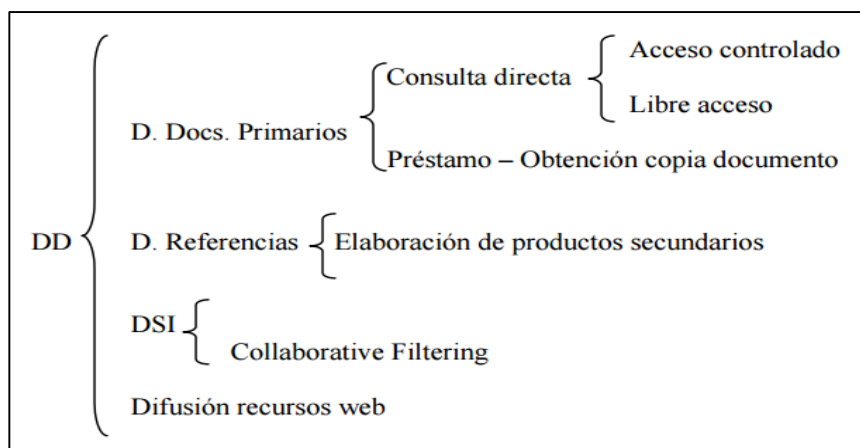


Figura 5: Difusión Documental. (Fuente: Castillo, L. (2004-2005))

La difusión documental es una difusión activa. Es aquella que implica iniciativa por parte del centro ofreciéndole, a los usuarios, productos documentales que juzga útiles, según el análisis que se haya hecho de sus necesidades, o señalándole sus problemas de información y ayudándole a resolverlos.

El centro deberá estudiar a sus usuarios con el objeto de averiguar qué tipo de información necesitan, en conjunto o por grupos, y debe estudiar también el sistema de difusión más adecuado para cada necesidad.

Dentro de la difusión documental se puede distinguir entre difusión de documentos primarios, la difusión de documentos secundarios o de referencia de documentos y la difusión selectiva de la información.

#### 1.4.4. Difusión de documentos primarios

##### a) Consulta directa en el centro

En acceso controlado. El usuario puede consultar los ficheros pero debe solicitar el documento que necesite a un empleado del centro. Se aplica a colecciones de documentos especiales por su contenido o soporte, obras raras o únicas, materiales débiles, materiales valiosos. Libre acceso. Los documentos se encuentran a disposición del usuario ordenados generalmente según criterios temáticos. Da más libertad al usuario pero disminuye el control de los fondos y requiere más espacio. También supone hacer frente a la complejidad que supone la ordenación de los fondos

##### b) Préstamos

Préstamos. El usuario obtiene el documento y puede utilizarlo durante un período de tiempo determinado. El documento debe ser devuelto. Obtención de copia del

documento. El usuario obtiene una copia del documento que pasa a ser de su propiedad, no debe devolverlo.

#### 1.4.5. Difusión de referencia de documentos

La mayoría de centros de documentación elabora algún tipo de publicación secundaria que contiene referencias de documentos y transmite a los usuarios para su conocimiento.

#### 1.4.6. Difusión selectiva de la información

Es un sistema de difusión “a la carta”, por el cual se ofrece a cada usuario las referencias de documentos correspondientes a sus temas de interés seleccionados a partir de todos los documentos recibidos durante un determinado período.

La difusión selectiva de información (DSI) ha sido tradicionalmente un problema no resuelto de Centros de Documentación, Centros de Información y Bibliotecas. Los perfiles de usuarios están, la información está, pero mantener un servicio de diseminación selectiva funcionando en forma eficiente consume significativos recursos organizacionales y es casi inevitable que en un procesamiento manual o semiautomático, que implica la administración de listas de usuarios, la elaboración de índices y boletines, la producción y compaginación de fotocopias, el mantenimiento de listas de direcciones de correo y teléfonos, la elaboración y etiquetados de sobres y el envío por correo o por fax, se cometan errores. No se puede pensar en atender requerimientos demasiado específicos con herramientas manuales o semiautomáticas. Para realizar un DSI automatizado se requiere algún programa informático que integre en un único concepto el acceso y la búsqueda en la base de datos, la agenda del sistema operativo y el correo electrónico. En los DSI automatizados las consultas a las bases de datos se realizan automáticamente y las respuestas que se obtienen se envían a cada usuario registrado en la lista de correos asociada al servicio. El mensaje de correo tendrá en forma anexa un documento con las referencias de aquellos documentos que se ajusten al perfil de información que ha solicitado o bien los textos íntegros de los documentos si el centro posee una base de datos fuente textuales.

#### 1.4.7. Difusión de recursos Web

Difusión de recursos en internet. Consiste en seleccionar, ordenar y describir las webs de interés para los usuarios del centro. Se distribuye mediante las redes internas del centro o intranet. (Ahora se empieza a aplicar el término de Arquitectura de la información a esta actividad). También puede consistir en

enviar, por correo electrónico, las URL de nuevas webs de interés. (Pérez Álvarez, 1998)

## **1.5. Restaurant**

Establecimiento público donde se sirven comidas y bebidas, mediante precio, para ser consumidas en el mismo local. (*Real Academia Española, 2017*).

### **1.5.1. Tipos de Restaurantes**

- Restaurante buffet o Tenedor libre. Es posible escoger una gran variedad de platos cocinados y dispuestos para el autoservicio. A veces se paga una cantidad fija y otras veces por cantidad consumida (por peso o tipos de platos). Surgido en los años setenta, es una forma rápida y sencilla de servir a grandes grupos de personas.
- Restaurante de comida rápida. Restaurantes informales donde se consume alimentos simples y de rápida preparación, como hamburguesas, patatas fritas, pizzas y pollo, entre otros.
- Restaurantes de alta cocina (restaurantes gourmet). Los alimentos son de gran calidad y se sirven a la mesa. El pedido es "a la carta" o se elige de un "menú", por lo que los alimentos se cocinan al momento. El costo depende del servicio y de la calidad de los platos que se consumen. Existen mozos o camareros, dirigidos por un Maitre. El servicio, la decoración, la ambientación, la comida y las bebidas se escogen cuidadosamente.
- Restaurantes temáticos. Se clasifican por el tipo de comida ofrecida. Los más comunes dependen del origen de la cocina, y los más populares en todo el mundo son: la cocina italiana y la cocina china, pero también la cocina mexicana, la cocina japonesa, la cocina española, la cocina francesa, la cocina peruana, la cocina colombiana, la cocina tailandesa y los restaurantes espectáculo.
- Restaurantes bares. Se sirve comida y bebida, y generalmente no se requiere consumir alimentos para poder pedir bebidas alcohólicas. Muy parecidos a las cantinas.

## **1.6. Definición de términos básicos**

### **1.6.1. Comensal**

Del latín co- 'co-' y mensa 'mesa'. Del cada una de las personas que comen en una misma mesa. Persona que vive a la mesa y expensas de otra, en cuya casa habita como familiar o dependiente (Barcia, R. (1894)).

### 1.6.2. Gastronomía

Proviene del griego "gaster" o "gastros", que quiere decir estómago y "gnomos". Es el estudio de la relación del hombre con su alimentación y su medio ambiente o entorno.

A menudo se piensa erróneamente que el término gastronomía únicamente tiene relación con el arte culinario y la cubertería en torno a una mesa. La gastronomía estudia varios componentes culturales tomando como eje central la comida. Se analizan actividades tales como degustar, transformar, historiar, narrar, descubrir, vincular, entender, conocer, contextualizar, experimentar e investigar los alimentos. Así pues, la gastronomía se debe entender como una actividad interdisciplinaria.

Si bien se observa se descubrirá que alrededor del alimento existe danza, teatro, pintura, escultura, literatura, arquitectura, música, en resumen, Bellas Artes. Pero también hay física, matemáticas, química, biología, geología, agronomía, es decir, ciencias exactas y naturales. Y además hay antropología, historia, letras, filosofía, psicología, sociología, por parte de las ciencias sociales. (DE ALVA, C. I. G (2012)).

### 1.6.3. Información científica

Conjunto de datos científicos que caracterizan una situación o materia, no conocidas por el sujeto, transmitidas a él por algún medio, que aumentan su conocimiento en algo y que pueden facilitar su actuación (Capurro, R. (2007)).

## CAPÍTULO II

### MATERIALES Y MÉTODOS

#### 2.1. Objetivos

##### 2.1.1. Objetivo general

Mejorar la difusión del servicio de recomendación de restaurantes en la ciudad de Tarapoto con el uso de un sistema basado en el modelo de recuperación vectorial.

##### 2.1.2. Objetivos específicos

- a) Conocer el servicio de recomendación para restaurantes en la ciudad de Tarapoto.
- b) Aplicar el modelo de recuperación vectorial, mediante la implementación de un buscador informático.
- c) Medir la influencia del buscador informático basado en el modelo de recuperación vectorial en el servicio de recomendación para restaurantes.

#### 2.2. Sistema de hipótesis

Una hipótesis es una suposición con la que el investigador trata de dar respuesta con lo conocido a lo desconocido en la búsqueda de nuevos conocimientos sobre la realidad.

En tal sentido se formuló las siguientes hipótesis:

**Hipótesis Alterna (Ha):** “Con la implementación de un Sistema basado en el Modelo de Recuperación Vectorial, se mejora el servicio de recomendación de restaurantes en la ciudad de Tarapoto”.

**Hipótesis Nula (Ho):** “Con la implementación de un Sistema basado en el Modelo de Recuperación Vectorial, no se mejora el servicio de recomendación de restaurantes en la ciudad de Tarapoto”.

##### 2.2.1. Sistema de variables

Por el lugar que ocupan en la hipótesis las variables pueden ser dependientes o independientes. En tal sentido se consideró las siguientes variables de estudio:

##### 2.2.2. Variable dependiente

Y: Servicio de recomendación de restaurantes.

##### Indicadores:

Y1: Nivel de difusión de información

Y2: Difusión bajo demanda

Y3: Difusión documental

Y4: Escala de difusión.



### 2.2.3. Variable independiente

X: Sistema basado en el modelo de recuperación Vectorial.

#### Indicadores:

X1: Precisión de los resultados de la búsqueda.

X2: Similitud Consulta-Respuesta.

X3: Tiempo de búsqueda.

Cabe mencionar que las variables se clasifican “por la naturaleza de sus valores en cualitativas y cuantitativas. Las cualitativas se presentan cuando sus estados representan cualidades del objeto estudiado, las cuantitativas son aquellas cuyos estados son cantidades continuas de un atributo que posee el objeto” (Taboada Neira, 2012).

Tabla 1

*Escala de medición de las variables.*

Variable	Indicador	Métricas	Escala de medición
Dependiente:  <b>Difusión de información de Restaurantes</b>	Y1: Nivel de difusión de información	<b>Unidades</b>	<b>Cuantitativa</b>
	Y2: Difusión bajo demanda	<b>Escala de linken</b>	Mala, Regular, Buena, Excelente
	Y3: difusión documental	<b>Escala de linken</b>	Mala, Regular, Buena, Excelente
	Y4: Escala de difusión		Mala, Regular, Buena, Excelente
Independiente:  <b>Sistema basado en el modelo de recuperación Vectorial</b>	X1: Similitud Consulta-Respuesta	<b>Escala de linken</b>	Mala, Regular, Buena, Excelente
	X2: navegación intuitiva	<b>Escala de linken</b>	Mala, Regular, Buena, Excelente
	X3: Sencillez de la interfaz	<b>Linker</b>	Mala, Regular, Buena, Excelente

Fuente: Elaborado por el investigador

### 2.3. Tipo de investigación

El estudio presentado corresponde a una investigación del tipo Aplicada - Transversal, porque los resultados obtenidos pretenden aplicarse en forma inmediata para ayudar a resolver la situación problemática planteada, analizando los resultados obtenidos en un mediano plazo.

### 2.4. Nivel de investigación

Experimental.- Porque se controló la variable independiente con el fin de mejorar el servicio de recomendación de restaurantes con el modelo de recuperación vectorial.

### 2.5. Diseño de investigación

Tabla 2

*Diseño Pre-Experimental*

<b>Estructura</b>	<b>Procedimientos</b>
<b>GE<sub>1</sub>: O<sub>1</sub> ---X--- O<sub>2</sub></b>	1) Precisión del grupo de estudio o muestra (Grupo experimental).
Donde:	
GE <sub>1</sub> : Grupo Experimental	2) Evaluación pre test al grupo de estudio (Grupo experimental).
O <sub>1</sub> : Evaluación pre test	3) Aplicación o tratamiento (al grupo experimental).
O <sub>2</sub> : Evaluación pos test	4) Evaluación pos test.
X: Sistema Informático	5) Comparación de los resultados de la evaluación pre test y evaluación pos test.
	<b>Utilidad:</b> Este diseño es útil para conocer los resultados de la aplicación que se ensaya.

Fuente: Elaborado por el investigador

### 2.6. Población y muestra

#### 2.6.1. Población

Según el reporte estadístico anual de arribos de turistas, tanto nacionales como extranjeros, que proporcionó el MINCETUR con respecto al distrito de Tarapoto en el año 2016, la población promedio mensual de turistas que arriban a la ciudad de Tarapoto fue de N = 33761 turistas.

#### 2.6.2. Muestra

Para el cálculo de la muestra se aplicó la fórmula de muestro de poblaciones finitas con un nivel de probabilidad de éxito del 95%.

$$n = \frac{N * Z\alpha^2 * p * q}{d^2 * (N - 1) + Z\alpha^2 * p * q}$$

Donde:

$$N = 33761$$

$$Z_{\alpha} = 1.962 \text{ (seguridad del 95\%)}$$

$p$  = proporción esperada (en este caso 90% = 0.9)

$q = 1-p$  (en este caso  $1 - 0.9 = 0.1$ )

$d$  = precisión (usaremos el 5%)

Al reemplazar los datos, en la fórmula tenemos:

$$n = \frac{33761 * 1.96^2 * 0.9 * 0.1}{0.05^2 * (33761 - 1) + 1.96^2 * 0.9 * 0.1}$$

$$n = \frac{6160,57224}{84,40+0,182476} = \frac{6160,57224}{84,582476} = 137,74$$

$$n = 138$$

$$137.74 = 138$$

La cual nos dio una muestra de valor de 138 turistas.

## 2.7. Técnicas e instrumentos de recolección de datos

Tabla 3

*Técnicas e instrumentos utilizados*

Técnica	Justificación	Instrumentos	Aplicado en...
1. Encuesta	1. Permite conocer la difusión de información de la biodiversidad de especies de flora y fauna.	1. Cuestionario.	1. A la muestra que es parte del estudio.
2. Registros	2. Proporciona la información del buscador informático.	2. Buscador Informático	2. Procesos que se desarrollan dentro del buscador.
3. Análisis de documentos	3. Para obtener la información de las fuentes secundarias referentes a temas de la investigación.	3. Fichas bibliográficas. Subrayado.	3. La bibliografía necesaria para desarrollar el marco teórico y la información complementaria.

Fuente: Elaborado por el investigador.

## 2.8. Técnicas de procedimiento y análisis de datos

### Técnicas de procedimiento

Para hallar que hay en los datos:

**La media aritmética o promedio (M):** Es el estadístico de tendencia central más significativo y corresponde variables de cualquier nivel de medición pero particularmente a las mediciones de intervalo y de razón.

$$M = \frac{X_1 + X_2 + X \dots}{N}$$

Dónde: M, media aritmética; X, frecuencia de un valor cualquiera de la variable y N, número total de los valores considerados.

**Desviación Estándar.-** Es una medida de dispersión de las variables cuantitativas.

**Programa de computador.-** Se utilizará el programa SPSS Statistics 22, versión para Windows 8.

**Tratamiento Estadístico.-** Los análisis estadísticos a emplear serán descriptivos e inferenciales, en la parte descriptiva se hará uso de tabla de frecuencia, gráficos y medidas como la media y desviación estándar. En la parte inferencial, para comparar los diferentes datos que se obtienen a través del instrumento elaborado para los fines de la investigación, también se usará la estadística T-student.

### Presentación de datos

- **Tablas:** Para un mejor entendimiento de los datos se presentarán en tablas distribuyéndolos en columnas y filas, agrupando los datos evitando duplicidades y superando el desorden. Servirán de ayuda visual que permitirán organizar los resultados de la investigación.
- **Figuras:** Importantes para expresar la tendencia de un hecho o fenómeno. Se usa figuras estadísticas como barras, líneas, áreas, etc.

## CAPÍTULO III

### RESULTADOS Y DISCUSIÓN

#### 3.1. Presentación de resultados

En nuestro estudio se llegó encuestar a 138 turistas de la ciudad de Tarapoto, llegándose a encontrar lo siguiente:

Tabla 4

*Resumen de los indicadores de la difusión, de información, antes de la implementación del buscador informático.*

Indicadores	Ítem	N	Porcentaje Marginal
DIFUSIÓN BAJO DEMANDA	Malo	22	16%
	Regular	101	73%
	Bueno	15	11%
	Excelente	0	0%
DIFUSIÓN DOCUMENTAL	Malo	25	18%
	Regular	95	69%
	Bueno	18	13%
	Excelente	0	0%
VÍA DE DIFUSIÓN UTILIZADA	Malo	35	25%
	Regular	85	62%
	Bueno	18	13%
	Excelente	0	0%
NIVEL DIFUSIÓN DE INFORMACIÓN	Malo	3	2%
	Regular	118	86%
	Bueno	17	12%
	Excelente	0	0%

Fuente: Elaborado por el investigador

En cuanto al indicador concerniente a la difusión bajo demanda, antes de la implementación del buscador informático, el 73% de los turistas lo considera como regular. (Ver Tabla 4).

Concerniente al indicador de la difusión documental, antes de la implementación del buscador informático, el 69% lo califica como regular. (Ver Tabla 4).

Además, el 62% de los turistas encuestados calificó como regular a la vía de difusión utilizada antes de la implementación del buscador informático. (Ver Tabla 4).

Por lo tanto en el nivel de difusión de información, se encontró que el 86% de los turistas lo calificaron como regular y solo el 12% como bueno, esto es antes de la implementación del buscador informático. (Ver Tabla 4 y Figura 6).

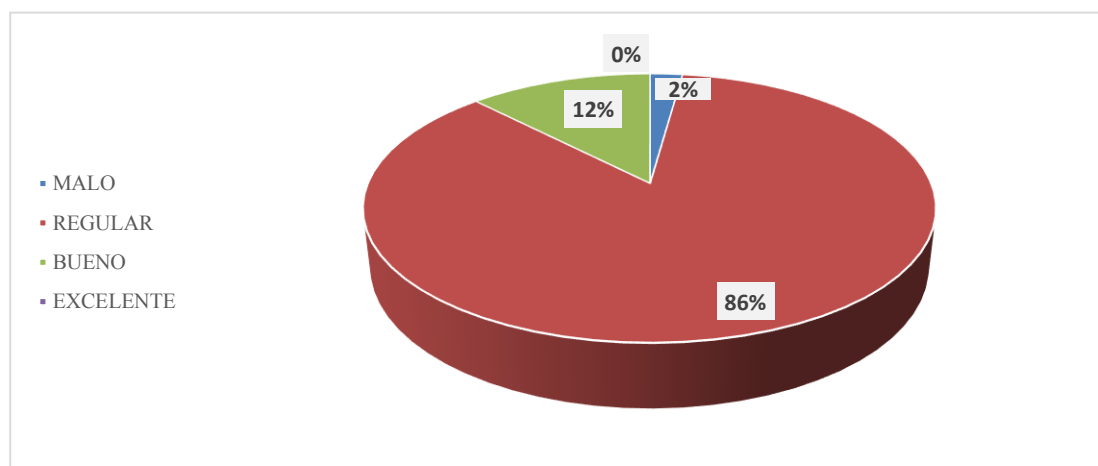


Figura 6: Nivel de difusión de información del Servicio de Recomendación de Restaurantes (Pre Test).  
(Fuente: Elaborado por el investigador)

Tabla 5

*Resumen de los indicadores de la difusión de información, después de la implementación del buscador informático*

Indicadores	Item	N	Porcentaje Marginal
DIFUSIÓN BAJO DEMANDA	Malo	0	0%
	Regular	2	1.4%
	Bueno	46	33.3%
	Excelente	90	65%
DIFUSIÓN DOCUMENTAL	Malo	0	0%
	Regular	2	1.45%
	Bueno	37	26.81%
	Excelente	99	71.74%
VÍA DE DIFUSIÓN UTILIZADA	Malo	0	0%
	Regular	1	1%
	Bueno	55	40%
	Excelente	82	59%

NIVEL DIFUSIÓN DE INFORMACIÓN	Malo	0	0%
	Regular	1	1%
	Bueno	18	13%
	Excelente	119	86%
SIMILITUD CONSULTA – RESPUESTA	Malo	0	0%
	Regular	6	4%
	Bueno	60	43%
	Excelente	72	52%
NAVEGACIÓN INTUITIVA	Malo	0	0%
	Regular	11	8%
	Bueno	51	37%
	Excelente	76	55%
SENCILLEZ DE LA INTERFAZ	Malo	0	0%
	Regular	6	4%
	Bueno	61	44%
	Excelente	71	51%

Fuente: Elaborado por el investigador.

Después de la implementación del buscador informático, se obtuvo el 65% de los turistas calificaron a la difusión bajo demanda como excelente. (Ver Tabla 5).

En cuanto a la difusión documental, para después de la implementación del buscador informático, se encontró el 71.74 % lo califica como excelente. (Ver Tabla 5).

En el caso de la vía de difusión, el 59% de los turistas lo calificó como excelente a después de la que se lleve a cabo la implementación del buscador informático. (Ver Tabla 5).

Por lo tanto en lo concerniente al nivel de difusión de información, se encontró que el 86% de los turistas lo calificaron como excelente y el 13% como bueno, esto después de realizar de la implementación del buscador informático. (Ver Tabla 5 y Figura 7).

Con respecto al buscador informático, entre el 52% de los turistas encuestados calificó como excelente a la similitud entre lo requerido y los resultados obtenidos al momento de realizar una búsqueda. (Ver Tabla 5).

En cuanto a la navegación intuitiva del buscador informático entre el 37% y el 100% de los turistas encuestados lo calificó como bueno y excelente. (Ver Tabla 5).

En cuanto a la sencillez de la interfaz del buscador informático entre el 55% y el de los turistas encuestados lo calificó como excelente. (Ver Tabla 5).

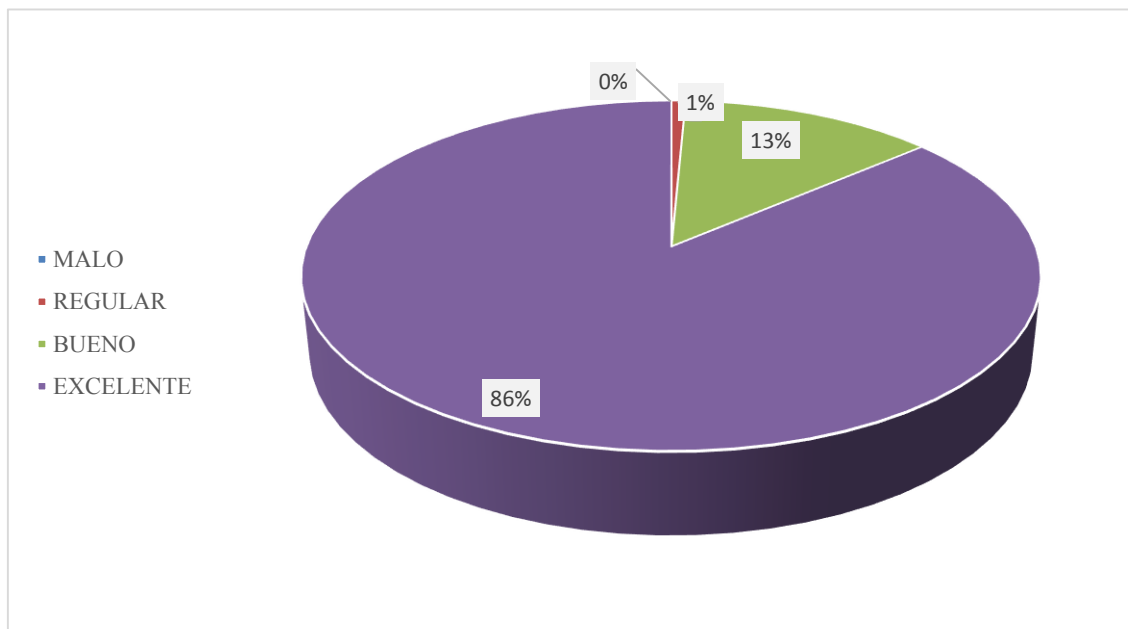


Figura 7: Nivel de difusión de información del Servicio de Recomendación de Restaurantes (Pos Test).  
(Fuente: Elaborado por el investigador)

### 3.2. Contrastación de hipótesis

#### 3.2.1. Definir la hipótesis

**Hipótesis Nula ( $H_0$ ):** “Con la implementación de un Sistema basado en el Modelo de Recuperación Vectorial, **no se mejora** el servicio de recomendación de restaurantes en la ciudad de Tarapoto”.

**Hipótesis Alternativa ( $H_a$ ):** “Con la implementación de un Sistema basado en el Modelo de Recuperación Vectorial, **se mejora** el servicio de recomendación de restaurantes en la ciudad de Tarapoto”.

Por lo tanto:

$$H_0 \Rightarrow U_A = U_B$$

$$H_a \Rightarrow U_A \neq U_B$$

Donde:

$$U_A = \text{Pre Test}$$

$$U_B = \text{Pos Test}$$

#### 3.2.2. Nivel de significancia

El nivel de significancia es:  $\alpha = 5\% = 0.05$ .



### 3.2.3. Estadístico de prueba

$$U_0 = \frac{\bar{X}_A - \bar{X}_B}{\sqrt{Sp^2 \left( \frac{1}{n_A} + \frac{1}{n_B} \right)}}$$

Donde:

$U_0$  = valor que representa a todo Pre y Pos test

$\bar{X}_A$  = media del Pre test (variable dependiente)

$\bar{X}_B$  = media del Pos test (variable dependiente)

$Sp^2$  = varianza de todo Pre y Pos Test

$n_A$  = total de encuestados el Pre test

$n_B$  = total de encuestados el Pos test

$S_A^2$  = varianza del Pre test

$S_B^2$  = varianza del Pos test

$$Sp^2 = \frac{S_A^2 * n_A + S_B^2 * n_B}{n_A + n_B - 2}$$

Entonces:

$$Sp^2 = \frac{4.11 * 138 + 3.75 * 138}{138 + 138 - 2}$$

$$Sp^2 = 3.96$$

$$U_0 = \frac{10.17 - 20.55}{\sqrt{3.96 \left( \frac{1}{138} + \frac{1}{138} \right)}}$$

$$U_0 = -43.35$$

Par ordenado  $\Rightarrow t_{(n_A+n_B-2, 1-\alpha/2)}$

### 3.2.4. Determinación de la región crítica

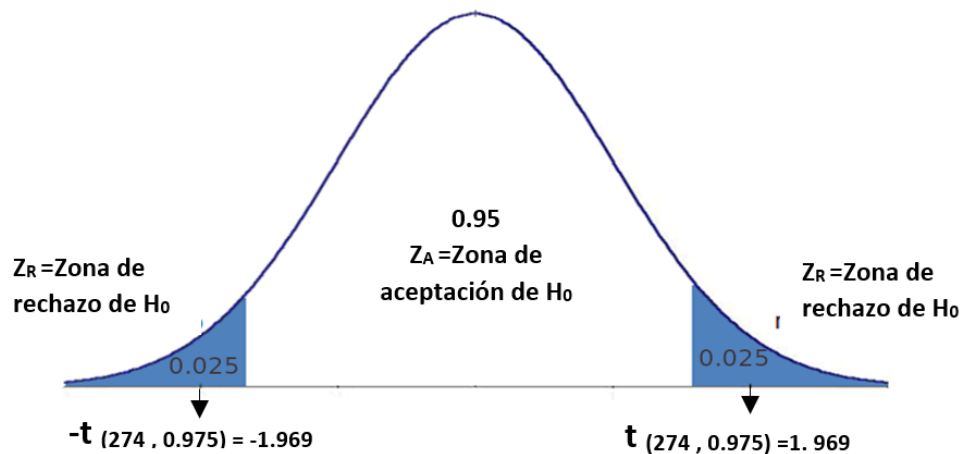


Figura 8: Región Crítica. (Fuente: Elaborado por el investigador)

Regla de decisión:

Aceptar  $H_0$  si  $U_0 \in Z_A$

Rechazar  $H_0$  si  $U_0 \notin Z_A$

### 3.2.5. Obtención del valor experimental

La utilización de la fórmula del estadístico de Prueba:

$$U_0 = \frac{10.17 - 20.55}{\sqrt{3.96\left(\frac{1}{138} + \frac{1}{138}\right)}}$$

$$U_0 = -43.35$$

Obtenemos que el valor de  $U_0$  es -43.35

### 3.2.6. Decisión

Se rechaza  $H_0$

### 3.2.7. Conclusión

Por lo tanto se puede afirmar que **se mejora** el servicio de recomendación de restaurantes en la ciudad de Tarapoto con la implementación de un Sistema basado en el Modelo de Recuperación Vectorial.

## La relación de la variable dependiente (Vx) sobre la variable independiente (Vy)

Donde se encuentra con los siguientes valores promediados del Pos test:

**Tabla 6**

*Valores Promediados del Pos Test*

X	Y
2.67	3.47
3	3.5
3.33	3.39
3.67	3.43
4	3.43

Fuente: Elaborado por el investigador

Dando como resultado el siguiente grafico de dispersión:

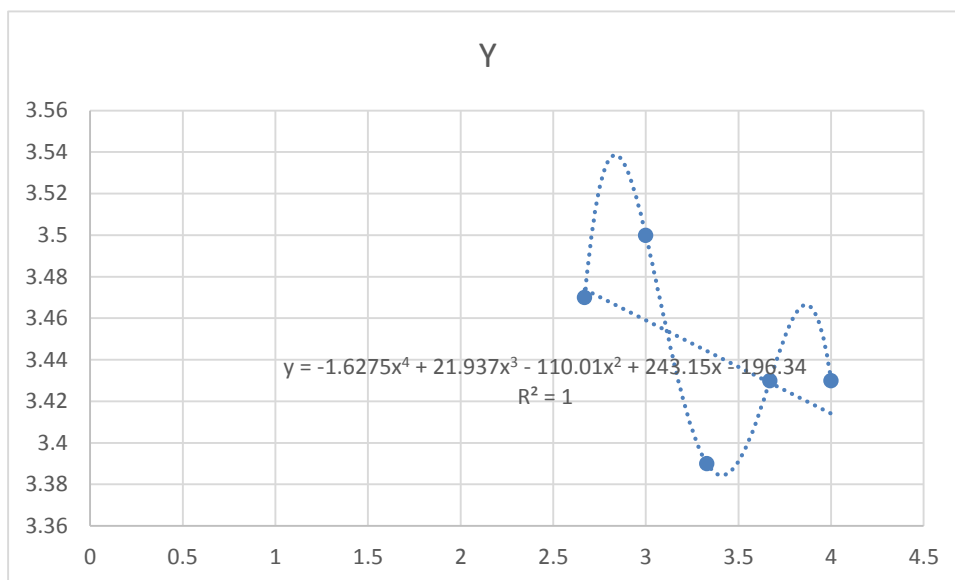


Figura 9: Grafica de dispersión. (Fuente: Elaborado por el investigador)

En el gráfico se observa la relación que existe de la Vx sobre la Vy, donde se aprecia que el coeficiente de determinación es  $R^2 = 1$  esta es una medida representativa de bondad del ajuste real, además que la línea de tendencia es polinómica de orden 4, la predicción del comportamiento de la recomendación de restaurantes en el mercado usando el sistema basado en el modelo de recuperación vectorial se explica en un 100%, lo cual manifiesta que las dos variables se ajustan perfectamente.

## CONCLUSIONES

Después del de realizar el respectivo análisis y discusión de los resultados obtenidos en el presente trabajo de investigación, se llegó a las siguientes conclusiones:

- Se realizó 138 encuestas a turistas, tanto en Pre y Pos Test, dando como resultado que el nivel de difusión en el Pre Test, el 86% de los turistas lo clasifico como Regular, 12% como bueno y un 2% como malo, por lo cual se concluye que el nivel de difusión es Regular, la cual muestra que la información que se difunde sobre el servicio de recomendación de restaurantes es mediocre; en cuanto a lo que se obtuvo en el Pos Test con la implementación del buscador, se encontró que los turistas lo clasifica como excelente con un 86% de aprobación, un 13% como bueno y 1% como regular, este concluye que el nivel de difusión con la implementación del buscador informático para el servicio de recomendación de restaurantes, logró un excelente aprobación.
- Se analizó los parámetros con relación al buscador informático, de lo cual los turistas encuentran que: la similitud consulta respuesta es excelente con un 52%, su navegación intuitiva es excelente con un 55% y la sencillez de interfaz es excelente con un 51% de aprobación.
- Se determinó que con la implementación de un buscador informático basado en el modelo de recuperación vectorial logró mejorar significativamente la difusión del servicio de recomendación de restaurantes en la ciudad de Tarapoto.
- Incorporar herramientas tecnológicas a medida de las necesidades de nuestra Región, las cuales genera un impacto positivo y en este caso mejoramos significativamente el servicio de recomendación de restaurantes de la ciudad de Tarapoto.
- La aplicación del Modelo de Recuperación Vectorial para el sistema informático de recomendación de restaurantes, nos ha permitido determinar que éste, es un buen modelo a usar, debido a que se adecua rápidamente al proceso de implementación de un sistema experto en diagnosis. ([www.sanmartincome.com](http://www.sanmartincome.com))

## **RECOMENDACIONES**

A continuación se formulan las siguientes recomendaciones:

- Realizar investigaciones que incorporen las tecnologías de información y comunicación que ayuden dentro de la cadena de valor turístico en nuestra región.
- Investigar nuevos modelos computacionales de aprendizaje de máquina que permitan implementar sistemas expertos en diagnóstico.
- Crear nuevas herramientas tecnológicas que apoyen al desarrollo sostenible de nuestra región.

## REFERENCIAS BIBLIOGRÁFICAS

- Amboage, E. S. (2015). El turismo 2.0. Un nuevo modelo de promoción turística. *REVISTA LUCIÉRNAGA-COMUNICACIÓN*, 3(6), 44-51.
- Barcia, R. (1894). Primer diccionario general etimológico de la lengua española (Vol. 5). Seix-Editor.
- Capurro, R. (2007). Epistemología y ciencia de la información. *Enlace*, 4(1), 11-29.
- Castillo, L. (2005). Difusión de la información. Obtenido de <http://www.uv.-es/macas>, 6.
- Castro Hidalgo, M., & Viale López, J. A. (Febrero de 2005). Sistema de búsqueda y comparación de documentos. Universidad Peruana de Ciencias Aplicadas.
- DE ALVA, C. I. G. (2012) Historia de la gastronomía.
- Definición y etimología de comensal (junio del 2017). Bogotá: E-Cultura Group. Obtenido de <https://definiciona.com/comensal/>
- Diccionario de la lengua española. (19 de mayo de 2017). Diccionario de la lengua española. Obtenido de Diccionario de la lengua española: <http://dle.rae.es/?id=M2v6jgO>
- EcuRed. (19 de mayo de 2017). EcuRed. Obtenido de EcuRed: [https://www.ecured.cu/Informaci%C3%B3n\\_Cient%C3%ADfica](https://www.ecured.cu/Informaci%C3%B3n_Cient%C3%ADfica)
- Gandara, J. M. G., Schlüter, R. G., Dieste, C. P., Lacanau, G. C., & Norrild, J. A. (2003). Gastronomía y turismo: una introducción. Ciet.
- Guamialamá, B. (2016). *Portal web para la promoción turística de la provincia de Carchi* (Bachelor's thesis).
- Joachims, T. (1998). Text categorization with support vector machines: Learning with many relevant features. *Machine learning: ECML-98*, 137-142.
- Lacanau, G. C., & Norrild, J. A. (2003). Gastronomía y turismo: cultura al plato. Centro de Investigaciones y Estudios Turísticos.
- Liu, T. Y. (2009). Learning to rank for information retrieval. *Foundations and Trends® in Information Retrieval*, 3(3), 225-331.

- Llopis, M. M. (1995). Historia de la gastronomía española. Val de Onsera.
- López Pérez, A. F. (2012). La gastronomía y su impacto turístico en la ciudad de Ambato provincia de Tungurahua en el periodo noviembre 2009 febrero 2010 (Bachelor's thesis).
- Marina, M. C. T. (2007). ¿ ES LA GASTRONOMÍA UN ARTE?. Hospitalidad ESDAI, (12).
- Martínez Méndez, F. J. (2002). Propuesta y desarrollo de un modelo para la evaluación de la recuperación de información en internet. Tesis Doctoral, Universidad de Murcia. España.
- Martínez Méndez, F. J. (2004). Recuperación de información: modelos, sistemas y evaluación. Murcia: Kiosko, 2004.
- Peñaranda Carvajal, J. M. (2013). Cultura Culinaria y Turística del Cantón Morona-Provincia de Morona Santiago, 2010(Bachelor's thesis).
- Pérez Ortiz, C. (2015). Análisis, diseño e implementación de una guía gastronómica para la administración y ubicación de restaurantes en entono Web.
- Robertson, S. (2004). Understanding inverse document frequency: on theoretical arguments for IDF. *Journal of documentation*, 60(5), 503-520.
- Suarez Lopez, L. G., & Mora Lozano, G. M. (2015). Diseño de una guía de gastronomía típica para Guayaquil (Bachelor's thesis, Espol).
- Subías, J. L. (2012). Lenguaje gastronómico: concepto y representación. *Tres notas. Disturbis*, (12), 0001-23.

## ANEXOS

### Anexo 01: Aplicación del pre y pos test.





Anexo 02: Instrumento de Pre-Test y Pos-Test**UNIVERSIDAD NACIONAL DE SAN MARTIN- TARAPOTO****FACULTAD INGENIERÍA DE SISTEMAS E INFORMÁTICA****FECHA:**

**Introducción:** La presente encuesta se realiza con la finalidad de ayudar en la investigación de cómo se puede mejorar el servicio de recomendación de restaurantes en la ciudad de Tarapoto. Con la información que se obtenga de esta encuesta, se diseñará un buscador para para recomendar los restaurantes según las necesidades.

**Estimad@ Turista:** agradecemos se sirva responder de manera veraz a esta encuesta pues los resultados ayudarán a desarrollar de una manera más precisa la presente investigación.

**Marque con una (X) en la alternativa correspondiente:**

*ITEMS PARA EVALUAR LA DIFUSION DE INFOMACION DE RETAURANTE DE LA CIUDAD DE TARAPOTO (PRE-TEST)*

**SOBRE LA DIFUSION DE INFOMACION DE RESTAURANTES****DIFUSION BAJO DEMANDA:**

1. La información sobre restaurantes de la ciudad de Tarapoto se encuentran sistematizados para su consulta de forma:
  - a. Malo
  - b. regular
  - c. bueno
  - d. excelente
2. La posibilidad de buscar información de restaurantes de la ciudad de Tarapoto utilizando diferentes criterios (ubicación precio tipo de comidas) es:
  - a. Malo
  - b. regular
  - c. bueno
  - d. excelente

**DIFUSION DOCUMENTAL:**

3. Modalidad (medios) que usa los restaurantes de Tarapoto para transmitir su información es:
  - a. Malo
  - b. regular
  - c. bueno
  - d. excelente
4. Cuando busca una información en particular las ofertas sugeridas por los restaurantes son:
  - a. Malo
  - b. regular
  - c. bueno
  - d. excelente

**VIA DE DIFUSION**

5. La posibilidad de obtener la información de restaurantes de Tarapoto:
  - a. Malo
  - b. regular
  - c. bueno
  - d. excelente
6. La disponibilidad de la información de restaurantes es:
  - a. Malo
  - b. regular
  - c. bueno
  - d. excelente

## UNIVERSIDAD NACIONAL DE SAN MARTIN- TARAPOTO

### FACULTAD INGENIERIA DE SISTEMAS E INFORMATICA

**FECHA:**

**Introducción:**

La presente encuesta se realiza con la finalidad de ayudar en la investigación de cómo se puede mejorar el servicio de recomendación de restaurantes en la ciudad de Tarapoto. Con la información que se obtenga de esta encuesta, se mejorará el buscador el cual recomienda los restaurantes según las necesidades.

**Estimad@ Turista:** agradecemos se sirva responder de manera veraz a esta encuesta pues los resultados ayudarán a desarrollar de una manera más precisa la presente investigación.

**Marque con una (X) en la alternativa correspondiente:**

*ITEMS PARA EVALUAR LA DIFUSION DE INFOMACION DE RETAURANTE DE LA CIUDAD DE TARAPOTO (POS-TEST)*

#### **SOBRE LA DIFUSION DE INFOMACION DE RESTAURANTES**

DIFUSION BAJO DEMANDA:

1. La información sobre restaurantes de la ciudad de Tarapoto se encuentran sistematizados para su consulta de forma:
  - a. Malo    b. regular    c. bueno    d. excelente
2. La posibilidad de buscar información de restaurantes de la ciudad de Tarapoto utilizando diferentes criterios (ubicación precio tipo de comidas) es:
  - a. Malo    b. regular    c. bueno    d. excelente

DIFUSION DOCUMENTAL:

3. Modalidad (medios) que usa los restaurantes de Tarapoto para transmitir su información es:
  - a. Malo    b. regular    c. bueno    d. excelente
4. Cuando busca una información en particular las ofertas sugeridas por los restaurantes son:
  - a. Malo    b. regular    c. bueno    d. excelente

VIA DE DIFUSION

5. La posibilidad de obtener la información de restaurantes de Tarapoto:
  - a. Malo    b. regular    c. bueno    d. excelente
6. La disponibilidad de la información de restaurantes es:
  - a. Malo    b. regular    c. bueno    d. excelente

**MODELO DE RECUPERACIÓN VECTORIAL:**

7. Cuando solicita información sobre restaurantes en el sistema de información ¿Cuál crees que es el nivel de similitud entre lo que usted requiere y los resultados obtenidos?
  - a. Malo
  - b. regular
  - c. bueno
  - d. excelente
8. Cuando usted usa el sistema de información ¿Qué calificación le daría a la navegación intuitiva del sistema de información?
  - a. Malo
  - b. regular
  - c. bueno
  - d. excelente
9. Cuando usted usa el sistema de información ¿Qué calificación le daría a la sencillez de la interfaz del sistema de información?
  - a. Malo
  - b. regular
  - c. bueno
  - d. excelente