



UNIVERSIDAD TÉCNICA DE COTOPAXI

DIRECCIÓN DE POSGRADO

MAESTRÍA EN SISTEMAS DE INFORMACIÓN

MODALIDAD: PROPUESTA METODOLÓGICA Y TECNOLÓGICA AVANZADA

Título:

Visualización de Información mediante mapeo auto-organizado en datos de producción científica de la Universidad Técnica de Cotopaxi.

Protocolo previo a la obtención del título de Magister en Sistemas de Información

Autor:

Rivera Chasiquiza Alex Geovanny

Tutor:

Cevallos Culqui Alex Santiago MSc.

LATACUNGA – ECUADOR

2020



APROBACIÓN DEL TUTOR

En mi calidad de Tutor del Trabajo de Titulación “Visualización de Información mediante mapeo auto-organizado en datos de producción científica de la Universidad Técnica de Cotopaxi” presentado por Rivera Chasiquiza Alex Geovanny, para optar por el título magíster en Sistemas de Información.

CERTIFICO

Que dicho trabajo de investigación ha sido revisado en todas sus partes y se considera que reúne los requisitos y méritos suficientes para ser sometido a la presentación para la valoración por parte del Tribunal de Lectores que se designe y su exposición y defensa pública.

Latacunga, septiembre, 16, 2020

.....
MSc. Cevallos Culqui Alex Santiago

CC.: 0502594427



APROBACIÓN TRIBUNAL

El trabajo de Titulación: Visualización de Información mediante mapeo auto-organizado en datos de producción científica de la Universidad Técnica de Cotopaxi, ha sido revisado, aprobado y autorizado su impresión y empastado, previo a la obtención del título de Magíster en Sistemas de Información; el presente trabajo reúne los requisitos de fondo y forma para que el estudiante pueda presentarse a la exposición y defensa.

Latacunga, septiembre, 16, 2020

.....
PhD. Gustavo Rodríguez Bárcenas
175700135-7
Presidente del tribunal

.....
PHD. Mayra Susana Alban Taipe
0502311988
Lector 2

.....
MSc. Jorge Bladimir Rubio Peñaherrera
0502222292
Lector 3



DEDICATORIA

Este trabajo dedico a mis padres Mario y Piedad por el apoyo incondicional recibido por su parte en continuar con mis estudios, a mis hermanas Gissela y Damaris por estar siempre pendientes a lo que me propongo, y a mi familia en general por estar presentes de una u otra manera durante este largo proceso.

Alex Rivera



AGRADECIMIENTO

Agradezco principalmente a mis padres y hermanas por estar presentes y apoyarme en cumplir mis sueños, a la Universidad Técnica de Cotopaxi por facilitar el proceso de aprendizaje competitivo y al MSc. Alex Cevallos que con su apoyo como tutor de investigación hace posible cumplir esta meta.

Alex Rivera



RESPONSABILIDAD DE AUTORÍA

Quien suscribe, declara que asume la autoría de los contenidos y los resultados obtenidos en el presente trabajo de titulación.

Latacunga, septiembre, 16, 2020

.....
Rivera Chasiquiza Alex Geovanny
C.I.: 05033497222



RENUNCIA DE DERECHOS

Quien suscribe, cede los derechos de autoría intelectual total y/o parcial del presente trabajo de titulación a la Universidad Técnica de Cotopaxi.

Latacunga, septiembre, 16, 2020

.....
Rivera Chasiquiza Alex Geovanny
C.I.: 0503349722



AVAL DEL VEEDOR

Quien suscribe, declara que el presente Trabajo de Titulación: Visualización de Información mediante mapeo auto-organizado en datos de producción científica de la Universidad Técnica de Cotopaxi contiene las correcciones a las observaciones realizadas por los lectores en sesión científica del tribunal.

Latacunga, septiembre, 20, 2020

.....
PhD. Gustavo Rodríguez Bárcenas
(1757001357)



UNIVERSIDAD TÉCNICA DE COTOPAXI DIRECCIÓN DE POSGRADO

MAESTRÍA EN SISTEMAS DE INFORMACIÓN

Título: Visualización de Información mediante mapeo auto-organizado en datos de producción científica de la Universidad Técnica de Cotopaxi

Autor: Rivera Chasiquiza Alex Geovanny

Tutor: Cevallos Culqui Alex Santiago MSc.

RESUMEN

La presente propuesta de investigación se enmarca en herramientas de visualización para tomar decisiones y clasificar los documentos científicos pertenecientes a docentes investigadores en la Universidad Técnica de Cotopaxi mediante mapeo auto-organizado, para lo cual fue necesario el diseño y desarrollo de una metodología que ayude a dar solución con la problemática presentada en el sistema ECUCIENCIA. En la actualidad las redes neuronales están siendo investigadas para el aprendizaje no supervisado y toma de decisiones, la metodología de mapeo auto-organizado (SOM) fue diseñada acorde a lo planteado por distintos investigadores y experiencias durante la presente indagación. La metodología SOM está estructurada con una serie de fases, desde la carga de datos en la cual debe existir un análisis y clasificación previa para el correcto entrenamiento de la red neuronal, lo que da como resultado una matriz de datos que posteriormente son agrupados y validados para su visualización. Los resultados obtenidos se los puede apreciar mediante mapas de calor, ya que estos expresan de forma interactiva y fiable para el usuario la visualización de los datos solicitados. Adicional es posible interactuar con la interfaz mediante filtros que muestran las líneas y sublíneas de investigación de acuerdo a la Universidad Técnica de Cotopaxi, los campos amplios, específicos y detallados según la CACES y UNESCO y el nivel de impacto en artículos de acuerdo a la categoría de la revista en la que fue subida en la cual según los resultados visualizados el investigador será recompensado.

PALABRAS CLAVE: mapeo auto-organizado; redes neuronales artificiales; inteligencia artificial; ECUCIENCIA; mapas de calor; aprendizaje no supervisado; toma de decisiones.



**UNIVERSIDAD TECNICA DE COTOPAXI
DIRECCION DE POSGRADO**

MAESTRIA EN SISTEMAS DE INFORMACIÓN

Title: INFORMATION VISUALIZATION THROUGH SELF-ORGANIZED MAPPING IN SCIENTIFIC PRODUCTION DATA OF THE TECHNICAL UNIVERSITY OF COTOPAXI.

Author: Rivera Chasiquiza Alex Geovanny
Tutor: Cevallos Culqui Alex Santiago MgC.

ABSTRACT

The present research proposal is framed in visualization tools to make decisions and classify the scientific documents belonging to professor -researchers at the Technical University of Cotopaxi through self-organized mapping, for which it was necessary to design and develop a methodology that helps to give solutions with the problematic presented in the ECUCIENCIA system. Currently neural networks are being researched for unsupervised learning and decision making, the methodology of self-organized mapping (SOM) was designed according to the raised by different researchers and experiences during the present inquiry. The SOM methodology is structured with a series of phases, from the data load in which there must be a prior analysis and classification for the correct training of the neural network, which gives as result a data matrix which later are grouped and validated for visualization. The results obtained can be seen through heat maps, as it they express in an interactive and reliable way for the user the viewing of the requested data. Additional it is possible to interact with the interface through filters that show lines and sublines of research according to the Technical University of Cotopaxi, the broad, specific and detailed fields according to CACES and UNESCO and the level of impact in articles according to the category of the journal in which it was uploaded in which according to the visualized results the researcher will be rewarded.

KEYWORD: self-organized mapping; artificial neural networks; artificial intelligence; ECUCIENCIA; heat maps; unsupervised learning; decision making.

Mayra Susana Córdor Martínez con cédula de identidad número:1723175350 Licenciado/a en: LICENCIADA EN CIENCIAS DE LA EDUCACIÓN MENCIÓN INGLÉS con número de registro de la SENESCYT: 1020-2018-2010749; **CERTIFICÓ** haber revisado y aprobado la traducción al idioma inglés del resumen del trabajo de investigación con el título: **“VISUALIZACIÓN DE INFORMACIÓN MEDIANTE MAPEO AUTO-ORGANIZADO EN DATOS DE PRODUCCIÓN CIENTÍFICA DE LA UNIVERSIDAD TÉCNICA DE COTOPAXI”** de: **Rivera Chasiquiza Alex Geovanny** aspirante a magister en Sistema de la información.

.....
Mayra Susana Córdor Martínez
1723175350

28 MAY 2020

Lic. Mayra Córdor
ENGLISH TEACHER

Latacunga, mayo 28 del 2020.

INDICE DE CONTENIDOS

APROBACIÓN DEL TUTOR	ii
APROBACIÓN TRIBUNAL	iii
DEDICATORIA	iv
AGRADECIMIENTO.....	v
RESPONSABILIDAD DE AUTORÍA.....	vi
RENUNCIA DE DERECHOS	vii
AVAL DEL VEEDOR.....	viii
RESUMEN	ix
INDICE DE CONTENIDOS	xi
ÍNDICE DE TABLAS	xiv
ÍNDICE DE FIGURAS	xvi
ÍNDICE DE ANEXOS.....	xvii
INTRODUCCIÓN	1
CAPÍTULO I	12
1. FUNDAMENTACIÓN TEÓRICA.....	12
1.1. ANTECEDENTES	12
1.2. FUNDAMENTACIÓN EPISTEMOLÓGICA.....	13
1.2.1. Sistemas informáticos	13
1.2.2. Sistemas informáticos en la web	13
1.2.3. Metodología ágil en el desarrollo de software	14
1.2.4. Modelo iterativo incremental	15
1.2.5. Ciclo de vida en el desarrollo de software.	16
1.2.6. Fases del modelo incremental	17
1.2.7. Lenguaje de programación.....	18

1.2.8.	Python	19
1.2.9.	Framework	20
1.2.10.	Framework web	20
1.2.11.	Desarrollo de aplicaciones web mediante la utilización de frameworks	21
1.2.12.	Django	22
1.2.13.	Red neuronal	23
1.2.14.	Red neuronal artificial	24
1.2.15.	Mapas auto-organizados	25
1.2.16.	Minería de datos	26
1.2.17.	Minería de texto	26
1.2.18.	Predicción	26
1.2.19.	Metodología	27
1.2.20.	Librería sompy	27
1.2.21.	Autoaprendizaje	28
1.2.22.	Toma de decisiones	28
1.2.23.	Manejo de entornos virtuales	28
1.2.24.	Áreas del conocimiento	29
1.2.25.	Base de datos	29
1.2.26.	PostgreSQL	30
1.3.	FUNDAMENTACIÓN DEL ESTADO DEL ARTE	30
1.4.	CONCLUSIONES CAPÍTULO I	34
	CAPÍTULO II	35
2.	PROPUESTA	35
2.1.	DIAGNÓSTICO	35
2.2.	MÉTODOS ESPECÍFICOS DE LA INVESTIGACIÓN	36

2.2.1.	Método de trabajo alineado a mapeo auto-organizado (SOM)	37
2.2.2.	Metodología de desarrollo	43
2.2.3.	Estrategias de implementación, comparativas y adopción tecnológica: 46	
2.2.4.	Modelos de análisis, diseño, implementación y prueba de sistemas de información.....	48
2.2.5.	Metodología de arquitectura de información y toma de decisiones:	52
2.2.6.	Estándares, ambientes y herramientas de software por utilizar:	55
2.3.	DISEÑO EXPERIMENTAL Y MÉTODO DE CRITERIO DE EXPERTOS.....	56
2.3.1.	Diseño experimental	56
2.3.2.	Método de criterio de expertos.....	58
2.4.	DESCRIPCIÓN METODOLÓGICA DE LA VALORIZACIÓN ECONÓMICA, TECNOLÓGICA, OPERACIONAL Y MEDIO AMBIENTAL DE LA PROPUESTA.	60
2.4.1.	Valoración económica:	60
2.4.2.	Valoración tecnológica:	60
2.4.3.	Valoración ambiental:.....	60
2.5.	CONCLUSIONES CAPITULO II	61
	CAPÍTULO III.....	62
3.	APLICACIÓN Y/O VALIDACION DE LA PROPUESTA	62
3.1.	RESULTADOS DE DIAGNÓSTICO DEL PROBLEMA	62
3.1.1.	Técnicas de investigación	63
3.2.	3.2. RESULTADOS DE LOS MÉTODOS ESPECÍFICOS.....	66
3.2.1.	MÉTODO DE TRABAJO ALINEADO A MAPEO AUTO- ORGANIZADO (SOM).....	67
3.2.2.	METODOLOGÍA DEL MODELO ITERATIVO INCREMENTAL 76	

3.3. RESULTADOS DEL DISEÑO EXPERIMENTAL	88
3.3.1. Método de criterio de expertos.....	88
3.4. RESULTADOS DE LA VALORIZACIÓN ECONÓMICA, TECNOLÓGICA, OPERACIONAL Y AMBIENTAL.....	95
3.4.1. Valoración económica:	95
3.4.2. Valoración tecnológica:	98
3.4.3. Valoración ambiental:.....	98
3.5. DISCUSIÓN DE LA APLICACIÓN Y VALIDACIÓN DE LA PROPUESTA	98
3.6. CONCLUSIONES CAPITULO III.....	100
CONCLUSIONES GENERALES.....	101
RECOMENDACIONES	102
BIBLIOGRAFÍA	103
ANEXOS.....	110

ÍNDICE DE TABLAS

Tabla 1. Sistema de Tareas	6
Tabla 2: Principios del Modelado Ágil	44
Tabla 3. Análisis de la Entrevista	64
Tabla 4. Contenido de involucrados.....	78
Tabla 5. Definiciones, acrónimos y abreviaturas.....	78
Tabla 6. Referencias.....	79
Tabla 7. Usuario Investigador.....	80
Tabla 8. Usuario.....	80
Tabla 9. RF01	81
Tabla 10. RF02	82
Tabla 11. RF03	82

Tabla 12. RF04	83
Tabla 13. RF05	83
Tabla 14. RNF01.....	84
Tabla 15. RNF02.....	84
Tabla 16. RNF03.....	84
Tabla 17. RNF04.....	85
Tabla 18. Confiabilidad.....	85
Tabla 19. Seguridad	86
Tabla 20. Resultados del procesamiento para la determinación del coeficiente de competencia de expertos	90
Tabla 21. Caracterización de los expertos que participaron en la consulta.....	90
Tabla 22. Definición de los criterios y escala para evaluar el método SOM.	93
Tabla 23. Estadísticos descriptivos de la valoración del cuestionario de evaluación.	94
Tabla 24. Alfa de Cronbach.....	94
Tabla 25. Gastos directos	96
Tabla 26. Gastos Indirectos	97
Tabla 27. Gastos Totales	98
Tabla 28. Historias de Usuario	113
Tabla 29. Historias de Usuario	113
Tabla 30. Ficha de Pruebas.....	118
Tabla 31. Informe de Pruebas	119
Tabla 32. Características de los sistemas operativos Linux y Windows.....	120
Tabla 33. Ventajas e inconvenientes de Linux como sistema operativo para servidores web	120
Tabla 34. Ventajas e inconvenientes de Windows como sistema operativo para servidores web	121
Tabla 35. Linux vs. Windows: tabla comparativa	121
Tabla 36. Requisitos mínimos del sistema operativo para implementar el módulo SOM.....	122

ÍNDICE DE FIGURAS

Figura 1. Metodología de gestión de proyectos.....	15
Figura 2: Modelo Iterativo Incremental.....	16
Figura 3: Ciclo de vida del desarrollo de software.....	17
Figura 4: Fases del modelo iterativo incremental.....	18
Figura 5: Metodología de Mapeo auto-organizado (SOM)	19
Figura 6: Lenguajes de programación más populares	20
Figura 7: Framework de aplicaciones web	21
Figura 8: Django: Esquema interno.....	22
Figura 9: Redes neuronales o el arte de imitar el cerebro.....	23
Figura 10: Red neuronal armada por científicos	24
Figura 11: Vigilancia Tecnológica de Ciclos Bioquímicos	26
Figura 12: Esquema de Base de Datos.....	30
Figura 13. Entrenamiento de las neuronas	40
Figura 14. Agrupamiento de los datos resultantes.....	40
Figura 15. Validación de la información.....	41
Figura 16. Visualización de información procesada.....	42
Figura 17. Modelo iterativo incremental.....	45
Figura 18. Elementos del Modelo de Análisis.....	48
Figura 19. Flujo de actividades de las dos fases de guía práctica.....	52
Figura 20. Mapa de navegación del módulo con el método SOM	54
Figura 21. Carga de Datos.....	68
Figura 22. Entrenamiento de las neuronas	69
Figura 23. Agrupamiento de los datos resultantes.....	70
Figura 24. Validación de la información.....	71
Figura 25. Validación de la información.....	71
Figura 26. Visualización de información procesada.....	72
Figura 27. Líneas de Investigación según la Universidad Técnica de Cotopaxi..	73
Figura 28. Sublíneas de Investigación según la Universidad Técnica de Cotopaxi	73
Figura 29. Campo Amplio según el CACES.....	75
Figura 30. Campo Específico según el CACES	75

Figura 31. Campo Detallado según el CACES.....	75
Figura 32. Etapas SOM con el modelo Iterativo Incremental	76
Figura 33. Años de experiencia en su respectiva área	91
Figura 34. Diagrama de Casos de Uso General.....	114
Figura 35. Diagrama de clases.....	114
Figura 36. Diagrama de Secuencia	115
Figura 37. Código View.py	115
Figura 38. Código som_listar.html	116
Figura 39. Mapas de calor con matplotlib.....	116
Figura 40. Mapa de calor con JavaScript	117
Figura 41. Listado de Líneas de Investigación de la Universidad Técnica de Cotopaxi	117

ÍNDICE DE ANEXOS

ANEXO I. Cuestionario de entrevista al coordinador de REDIC.....	111
ANEXO II. Historias de Usuario	113
ANEXO III. Diagrama de casos de uso general	114
ANEXO IV. Diagrama de clases y Diagrama de secuencia	114
ANEXO V. Resultados de la implementación	115
ANEXO VI. Ficha e Informe de Pruebas.	118
ANEXO VII. Servidor de alojamiento del módulo SOM	120

INTRODUCCIÓN

El mapeo auto-organizado (SOM) ha sido un tema importante dentro de la investigación para el análisis y visualización de datos, además, utilizado en la creación de modelos predictivos. Los SOM fueron presentados por T. Kohonen en 1982 como un modelo de red, basado en ciertas evidencias descubiertas a nivel cerebral. Este tipo de red posee un aprendizaje no supervisado competitivo [1]. Una red auto-organizada debe diferenciar los datos de entrada para incorporarlos a la estructura interna del mismo, es decir autoorganizarse en función a los datos. El aprendizaje competitivo consiste en que las neuronas compitan entre ellas para realizar una tarea o activarse, pretendiendo que solo una neurona o grupo vecinas se active a la salida.

Según [1] expresa acerca de los SOM, siendo parte de Redes Neuronales Artificiales han sido investigadas desde la década de 1980, y se han implementado en lenguajes de programación como C, Fortran, R y Python. Python al ser un lenguaje de alto nivel, eficiente y ampliamente utilizado en el campo del aprendizaje autónomo durante años posee una gran gama de paquetes relacionados con SOM para la construcción y visualización de modelos. Según [2] el paquete (librería) SOMPY, es capaz de realizar funcionalidades más allá de la construcción y visualización de modelos, tales como entrenamiento por lotes, que es más rápido que el entrenamiento en línea, además realiza diferentes métodos para la aproximación y predicción.

Este aprendizaje pretende categorizar los datos introducidos en la red, clasificar valores en la misma categoría y, por tanto, activar la misma neurona de salida [1]. Según [3], la Inteligencia Artificial, ha sido una herramienta de soporte importante en diferentes ámbitos de actividad, permitiendo la agregación de conocimientos, la optimización de procesos y la aplicación de metodologías capaces de resolver problemas reales complejos. A pesar de centrarse en una amplia gama de métricas exitosas, el enfoque de la Red Neuronal Artificial (RNA), una técnica similar al sistema nervioso central ha ganado notoriedad en la actualidad.

Al presente la Inteligencia Artificial (IA) está siendo investigada con más ímpetu que en décadas pasadas, esto se debe a que se ha visto un camino fiable para las

industrias que ofrecen productos a clientes, es así que se está incursionando en las diferentes áreas que nos ofrece, una de las principales son las redes neuronales que son la base para la toma de decisiones y el autoaprendizaje de sistemas.

Las técnicas de IA han sido aplicadas al estudio y resolución de problemas complejos, alcanzando efectividad en los resultados, por tal motivo ha sido adaptada tanto a diagnóstico de enfermedades, venta de calzado, marketing, entre muchos más ejemplos que se podrían mencionar. Las RNA han demostrado resultados acertados y fiables partiendo de bases de datos. En la referencia [4] se explica que las redes neuronales no consideran la incertidumbre, actúan como una caja negra, en el sentido de que no es posible saber cómo se ha llegado a los resultados obtenidos y los nodos intermedios no pueden ser interpretados.

En [5] se explica que ThemeRiver proporciona a los usuarios una vista macro de los cambios temáticos en un conjunto de documentos en una dimensión en serie. Muchos sistemas de visualización de información representan cada documento o grupo de documentos con un glifo o icono, en función de diversos atributos del documento. Pocos sistemas representan cambios a lo largo del tiempo, pero en los que lo hacen, los cambios de horas extras generalmente se representan mediante una secuencia de imágenes o "cuadros". Cada cuadro muestra solo los documentos relacionados con un intervalo de tiempo específico.

Además [6] explica que ThemeRiver es un sistema prototipo que visualiza variaciones temáticas a lo largo del tiempo dentro de una gran colección de documentos. En la visualización de información exploratoria se intenta presentar la información para que los usuarios puedan distinguir fácilmente los patrones. Siendo estos los que revelan tendencias, relaciones, anomalías y estructuras de datos para ayudar a los usuarios.

De acuerdo a [7], las redes bayesianas calculan la probabilidad de que un caso con cierto atributo pertenezca a una categoría. En medicina se han utilizado para predecir recaída de enfermedades o riesgo de cáncer. También poseen ventajas frente a las redes neuronales, el razonamiento basado en casos para distinguir al ejecutar el diagnóstico en enfermedades. Teniendo ventaja al razonamiento humano al diagnosticar las mismas.

En la actualidad la mayoría de investigaciones están encaminadas en la IA, siendo esto importante ya que permite en el área de informática desarrollar sistemas complejos. Las tecnologías de información y comunicación actualmente son herramientas de suma importancia para realizar investigaciones, debido a que permiten el acceso, producción, tratamiento y comunicación de información presentados en diferentes contextos. Implícitos en las TIC tenemos sublíneas, como la inteligencia artificial y la inteligencia de negocios para toma de decisiones. De acuerdo a [8] la IA es la combinación de algoritmos planteados con el propósito de crear maquinas que presenten las mismas capacidades que el ser humano. Según expresa que la inteligencia de negocios es un conjunto de estrategias que surgen de la información generada por herramientas, las cuales tienen como objetivo analizar una gran cantidad de datos que le dan al negocio la oportunidad de tomar decisiones con datos reales.

La presente investigación, se enfoca en el uso de redes neuronales (una rama de la inteligencia artificial) y algoritmos genéticos para el auto aprendizaje de dichas neuronas, es decir la predicción de resultados mediante la toma de decisiones. Siendo este tema uno de mucha relevancia en la actualidad, dando firmeza para que muchos investigadores quieran inclinarse por esta línea y sublínea de investigación. De acuerdo a [9] las Redes Neuronales Artificiales (RNA) son dirigidos a partir de los datos, es decir, son capaces de encontrar relaciones (patrones) de forma inductiva por medio de los algoritmos de aprendizaje basados en los datos existentes. Por su parte acerca de los algoritmos genéticos [10] nos dice que es un método de búsqueda que imita la teoría de la evolución biológica de Darwin para la resolución de problemas. A partir de esta breve explicación según [4] en la actualidad existe gran cúmulo de información que aparecen en libros, revistas, bases de datos y otros soportes, sin embargo, al encontrarse tan dispersa, resulta muy engorrosa la búsqueda e identificación de diferentes temas que son importantes para continuar con la investigación. En este sentido se han desarrollado herramientas de análisis que agrupan la información por temática y otras que permiten visualizar este conocimiento. Estas herramientas se conocen con el nombre de Escalamiento Multidimensional y Mapas Autoorganizados (SOM) o algoritmos de Kohonen.

Según [11] el auge y generalización en el estudio de potentes técnicas Cienciométricas, es el pilar fundamental en el que se basan los actuales procesos de evaluación de instituciones, programas docentes e investigadores. Un claro ejemplo son los SOM para la evaluación de la investigación de tesis doctorales: el caso de la Didáctica de las Ciencias Sociales en España publicada por la Revista Electrónica de Investigación y Evaluación Educativa. De acuerdo con [12], en el contexto actual, nos encontramos con una colección de problemas, que mediante el uso de algoritmos de aprendizaje se puede llegar a obtener buenas aproximaciones en un tiempo mucho más razonable. Esto ocurre con el problema del viajante de comercio, que usando una aproximación mediante una red de Kohonen adaptada al problema, puede conseguir una mejora notable en el tiempo de ejecución.

Según [1] expresa que los profesionales que necesitan manejar grandes volúmenes de información por las características de su trabajo saben que uno de los retos actuales es conocer y dominar el uso de herramientas que les permitan procesar grandes almacenes de datos. A esta área de trabajo se la conoce como minería de datos y posteriormente minería de textos. Muchos problemas de las disciplinas relacionadas con los estudios métricos de la información requieren de la aplicación de estas técnicas.

Las redes neuronales artificiales (RNA) han probado ser de gran utilidad para resolver problemas de minería de datos y texto. Estas han sido útiles particularmente en la organización de la información, el descubrimiento de conocimiento y la visualización de información. Esta última es entendida como el “proceso de interiorización del conocimiento mediante la percepción de información”. La visualización de información, según la definición anterior, interviene en el paso de datos a información y en la posibilidad de la construcción del conocimiento, al revelar los patrones que subyacen a los datos [1].

Actualmente las investigaciones se enfocan al manejo de la Big Data, debido a que permite predecir las necesidades y deseos de clientes, de cierta forma acelerando el crecimiento de la industria. La capacidad de predecir las tendencias en función a las necesidades del consumidor significa una ventaja considerable con la competencia.

A pesar de los beneficios que el uso de SOM propone, la mayoría de los sistemas aún no se encuentran preparados para explotar la información existente en el entorno, principalmente por la carencia de herramientas que faciliten el análisis de la data. En este contexto, investigadores de la Universidad Técnica de Cotopaxi se encuentran desarrollando un sistema el cual tiene por nombre ECUCIENCIA, el mismo, estudia la colaboración de artículos, libros y ponencias científicas de docentes-investigadores en la institución.

La situación es que existe una difícil clasificación por áreas del conocimiento en los datos de producción científica de docentes investigadores de la Universidad Técnica de Cotopaxi ya que se requiere automatizar el proceso conforme a sus líneas y sublíneas de investigación y de acuerdo a los campos amplios, específicos y detallados según la CACES y UNESCO. Mediante estas definiciones planteadas visualizar los datos mediante mapas de calor, representando esto una mejor interacción con los usuarios del sistema.

Formulación del problema: ¿Cómo visualizar los datos de producción científica de los docentes-investigadores en la Universidad Técnica de Cotopaxi donde existe difícil clasificación por áreas del conocimiento según sus líneas y sublíneas y de acuerdo al CACES y UNESCO según sus campos amplios, específicos y detallados?

EL **Objetivo General** planteado es establecer herramientas de visualización para la toma de decisiones y clasificación de documentos científicos pertenecientes a docentes investigadores en la Universidad Técnica de Cotopaxi a través del método SOM (Mapeo Auto-Organizados) con entrenamiento de redes neuronales.

Por su parte los **Objetivos Específicos** propuestos para la investigación son:

- Analizar los estudios realizados por diferentes investigaciones referentes a SOM y redes neuronales artificiales interpretando sus resultados publicados en revistas científicas para implementarlos en el Sistema ECUCIENCIA.
- Implementar redes neuronales artificiales para la toma de decisiones y Machine Learning en la clasificación por áreas del conocimiento según las líneas de investigación en la Universidad Técnica de Cotopaxi, CACES y UNESCO.

- Visualizar la información de forma clara y concisa mediante el método SOM en la clasificación por áreas del conocimiento según las líneas de investigación en la Universidad Técnica de Cotopaxi, CACES y UNESCO.

A continuación, se detallan las tareas necesarias para la ejecución de la propuesta de investigación.

Tabla 1. Sistema de Tareas

Objetivo	Actividad (tareas)
1. Analizar los estudios realizados por diferentes investigaciones referentes a SOM y redes neuronales artificiales interpretando sus resultados publicados en revistas científicas para implementarlos en el Sistema ECUCIENCIA.	1. Indagar en revistas Científicas e investigaciones referentes al tema.
	2. Analizar resultados obtenidos en investigaciones referentes a SOM.
	3. Relacionar con el mapeo auto-organizado.
2. Implementar redes neuronales artificiales para la toma de decisiones y Machine Learning en la clasificación por áreas del conocimiento según las líneas de investigación en la Universidad Técnica de Cotopaxi, CACES y UNESCO.	1. Analizar algoritmos genéticos para la toma de decisiones.
	2. Preparación de la data de entrada para el aprendizaje.
	3. Implementar redes neuronales y algoritmos genéticos en el sistema web a desarrollarse.
	1. Diseñar un prototipo inicial.

3. Visualizar la información de forma clara y concisa mediante el método SOM en la clasificación por áreas del conocimiento según las líneas de investigación en la Universidad Técnica de Cotopaxi, CACES y UNESCO.	2. Visualizar la información de forma ordenada y concisa.
	3. Asociar la visualización al sistema ECUCIENCIA.

Elaborado por: Alex Geovanny Rivera Chasiquiza

Se determina la **justificación** de la propuesta de investigación, por varios parámetros descritos en el siguiente apartado. La implementación de Inteligencia Artificial (IA), Redes Neuronales Artificiales (RNA) y Mapas Auto-organizados (SOM) en la actualidad se encuentra en constante evolución ya que han abierto nuevas rutas para el manejo y visualización de grandes cantidades de datos en sistemas alojados en la web principalmente. Los SOM se encargan de entrenar y agrupar por similitud una data de entrada sin supervisión para obtener a la salida una proyección de colores que serán visualizados por los usuarios, aportando considerablemente en la toma de decisiones que comúnmente serían complejas a simple vista.

La presente propuesta de investigación es visionaria, ya que se enmarca en la toma de decisiones no supervisadas en el sistema Web ECUCIENCIA que se encuentra en constante desarrollo en la Universidad Técnica de Cotopaxi. La visualización de información con SOM se aplicará a las Líneas y Sublíneas de Investigación de acuerdo a la Universidad Técnica de Cotopaxi, y a los Campos Amplios, Específicos y Detallados según la CACES y UNESCO. La data a ser manipulada para el proceso de SOM será tomada de los artículos, libros y ponencias científicas publicadas por docentes investigadores de la Universidad Técnica de Cotopaxi en revistas científicas, las cuales serán analizadas por la calidad de las mismas.

El impacto de la implementación de SOM en el sistema ECUCIENCIA es muy beneficioso para los docentes investigadores de la Universidad Técnica de Cotopaxi ya que con su uso permitirá que los procesos de producción científica producidos en la institución sean evaluados y categorizados de acuerdo a la conmovición que los mismos generen en la sociedad según los lineamientos en la Universidad, CACES y UNESCO otorgando un mejor control de la planta docente en la institución, además que ayudara a propagar el uso de tecnologías actualizadas que se encuentran en producción en distintas partes del mundo, siendo esto una gran ventaja para la Universidad Técnica de Cotopaxi y sus docentes investigadores.

Existen muchas **metodologías** aplicables para todo lo que abarca la Inteligencia Artificial (IA), es así que a lo largo de la investigación se vio pertinente la creación de una metodología que permita abarcar los procesos desde la concepción y carga de datos, continuando por su posterior entrenamiento y agrupamiento, para finalmente validar y visualizar los resultados. Esta metodología será de gran ayuda el manejo de grandes cantidades de datos, debido al potencial mostrado en el modelado y control de procesos enfocados a neuronas artificiales (ANN) e Inteligencia Artificial como tal. Las metodologías de aprendizaje no supervisado aún están en proceso de prueba, aunque la metodología SOM ha mostrado un aumento constante en investigaciones enfocadas en tecnologías futuras debido a su solidez.

La Método de trabajo alineado al Mapeo auto-organizado (SOM) será relevante durante el desarrollo y ejecución de la propuesta ya que se pretende analizar, clasificar y visualizar información científica de intelectuales a partir de las líneas y sublíneas de investigación acorde a la Universidad Técnica de Cotopaxi, CACES y UNESCO. Al ser una RNA, está basada en varias etapas, las cuales de acuerdo a lo analizado en el método de trabajo alineado a SOM son la carga de datos, entrenamiento, agrupamiento, pruebas, validación y visualización de los resultados del proceso como se muestra en la Figura 1 a continuación, permitiendo validar el objetivo principal de la investigación y obteniendo los resultados esperados.

Adicional a la metodología propuesta, será necesario utilizar la metodología del enfoque cuali-cuantitativo, debido a su estructura en la recolección y análisis de

datos, además que permitirá obtener resultados óptimos y fiables de la propuesta de investigación planteada para la visualización de información de docentes investigadores de la Universidad Técnica de Cotopaxi mediante mapeo autoorganizados. Se logra llegar a la información necesaria utilizando técnicas que esta metodología propone, como son la entrevista, observación y encuestas, que son importantes para su posterior análisis y posteriormente para la visualización propuesta.

La propuesta de investigación está enfocada al análisis de requerimientos planteados por el usuario encargado del sistema web ECUCIENCIA, por lo cual resulta prudente tener el primer acercamiento real a lo propuesto mediante la investigación explicativa, esta técnica muestra los procesos realizados en la producción científica en la Universidad Técnica de Cotopaxi. Además, se aplicará la investigación descriptiva como siguiente paso para obtener la situación real de dicha producción, ayudando a identificar la problemática suscitada en el sitio web. Debido a la existencia de investigaciones realizadas en distintas partes del planeta, se encuentra racional utilizar la metodología expresada, siendo de vital apoyo para identificar los requerimientos necesarios para dar solución a lo que viene aconteciendo con los procesos de producción científica en la Universidad Técnica de Cotopaxi.

Los beneficiarios directos serán los docentes investigadores de la Universidad Técnica de Cotopaxi, la cual consta de 470 docentes, los mismos que serán tomados en cuenta como usuarios del sitio web ECUCIENCIA debido a su importancia en el proceso de producción científica. Basados en el total de población recopilados se considera no calcular una muestra, siendo tomada toda la población para la investigación de la propuesta planteada referente a la visualización de información de docentes investigadores de la Universidad Técnica de Cotopaxi mediante mapeo auto-organizados.

Los **métodos teóricos** de acuerdo a [13] permiten descubrir en el objeto de investigación las relaciones esenciales y las cualidades fundamentales, no detectables de manera sensorial, es por ello que se apoya básicamente en los

procesos de abstracción, análisis, síntesis, inducción y deducción. Debido a lo mencionado los métodos teóricos a utilizarse en la investigación serán:

- **Método Deductivo:** Consiste en la forma de razonamiento que permite pasar de un conocimiento general a uno particular, es decir una estrategia de razonamiento empleada para obtener la conclusión que se encuentra inmersa en las premisas que este método interpreta. El Método Deductivo ayuda a deducir los datos verídicos en la investigación de la propuesta.
- **Método Inductivo:** Por su parte el método inductivo suele decirse es lo opuesto del método deductivo, es decir obtiene las conclusiones a partir de premisas particulares. Siendo el método científico más usual debido a los pasos esenciales que ofrece como la observación, clasificación y estudio, derivación y generalización de los hechos en particulares. El Método Inductivo permite el descubrimiento de un hecho desconocido a partir de los hechos conocidos para llegar a una solución fiable del problema expuesto.
- **Método histórico-lógico:** Este método se caracteriza por la trayectoria de los fenómenos a través de su evolución y desarrollo histórico, siendo la lógica la encargada de reproducir la esencia del objeto de estudio. Es decir que lo lógico como lo histórico se complementan mutuamente para descubrir las causas de los fenómenos y llegar a una conclusión o resultado de la problemática.

En este sentido este método será fundamental en la investigación planteada debido a que permitirá interpretar de mejor manera las fuentes bibliográficas y sus contenidos para dar solución al problema planteado en la visualización de información de docentes investigadores en la Universidad Técnica de Cotopaxi mediante mapeo auto-organizados.

Según [13] dice que los métodos **empíricos** aportan al proceso de investigación con resultados de la experiencia a través de procedimientos prácticos con diversos medios de estudio. Los métodos empíricos a utilizarse son los siguientes:

- **Observación:** Esta técnica se la pone en práctica, poniendo atención a nuestro entorno de acuerdo a la realidad en la problemática para recoger datos referentes a la investigación para su análisis e interpretación. Al ser una técnica primaria, permite recolectar datos desde el lugar donde ocurre la problemática

o hechos. En relación a la temática propuesta en esta investigación, se realiza en el sitio web junto con su contenido, es decir el lugar donde se procesa los datos de producción científicos de la Universidad Técnica de Cotopaxi.

- **Entrevista:** Está técnica hace referencia a la recolección oral de datos al o los encargados en el lugar que presenta la problemática de forma directa y concisa. Esta técnica será aplicada al PhD Gustavo Rodríguez encargado del sitio web ECUCIENCIA, en el cual se pretende implementar la investigación para la clasificación y visualización de información científica de docentes investigadores de la Universidad Técnica de Cotopaxi.

El **Método de trabajo alineado a mapeo auto-organizado (SOM)** describe que bajo los recursos encontrados y partiendo de investigaciones relacionadas a redes neuronales y aprendizaje no supervisado en IA, se desarrolló un método de trabajo acorde a las necesidades en la investigación para la resolución del problema encontrado en el proceso de producción científica en la Universidad Técnica de Cotopaxi. Este método de trabajo está planteado para la carga de data, entrenamiento, agrupamiento, validación y visualización de los resultados, permitiendo el proceso sea interactivo y confiable para su análisis.

CAPÍTULO I.

1. FUNDAMENTACIÓN TEÓRICA

1.1. ANTECEDENTES

La clasificación de datos siempre ha sido un tema de estudio e investigación a nivel mundial y sobre todo la clasificación de documentación científica, esto se debe principalmente al considerable incremento de información en la actualidad y los métodos tradicionales que se emplean. A nivel internacional y nacional existen investigaciones realizadas en diferentes áreas relacionadas con la inteligencia artificial (IA), redes neuronales artificiales (RNA), mapeo auto-organizado (SOM), algoritmos genéticos, toma de decisiones y aprendizaje no supervisado. Por lo que esta investigación se nutre de fuentes reales y confiables tomadas como antecedentes teóricos.

Durante la indagación, se observa que en Brasil se han elaborado la mayor parte de investigaciones relacionadas a la Inteligencia Artificial, lo que conlleva implementarla para el uso común, como es el caso de la toma de decisiones, el aprendizaje supervisado o en el caso de la propuesta planteado acerca del mapeo auto-organizado de datos de producción científica de docentes-investigadores de la Universidad Técnica de Cotopaxi. También se observan investigaciones realizadas en Europa, siendo casos en Turquía, Japón, España, poniendo como ejemplo a ciertos países que se encuentran en esta línea de investigación.

Algo que también es necesario mencionar, es el hecho que en nuestro país existen muy pocas evidencias de investigaciones relacionadas a Inteligencia Artificial o por lo menos no se encuentran publicados, se entiende que en el Ecuador se está dando los primeros pasos en lo que a IA se refiere. También, por lo mencionado se tomará

como referencia en la investigación planteada, los estudios realizados en Europa y Asia principalmente, ya que en centros de investigaciones en dichos lugares son donde más a fondo se han realizado las indagaciones, dando un punto de partida para explotar estos conocimientos y esta ciencia que esta día a día enmarcándose con más fuerza en la cotidianidad, principalmente de sistemas educativos y empresas orientadas a los sistemas de información.

1.2. FUNDAMENTACIÓN EPISTEMOLÓGICA

1.2.1. Sistemas informáticos

En el 2018, según [14] dice que los sistemas de información, constituyen uno de los principales ámbitos de estudio en el área de organización de empresas. El entorno donde las compañías desarrollan sus actividades, se vuelve cada vez más complejo. Según [15], un sistema de información queda definido como: conjunto formal de procesos que, operando sobre una colección de datos estructurada de acuerdo a las necesidades de la empresa, recopila, elabora y distribuyen selectivamente la información necesaria para la operación de dicha empresa y para las actividades de dirección y control correspondientes, apoyando, al menos en parte, los procesos de toma de decisiones necesarios para desempeñar funciones de negocio de la empresa de acuerdo con su estrategia.

De acuerdo a [16] los desempeños de los Sistemas Informáticos han ocupado un papel muy importante, principalmente en la toma de decisiones y trabajos laborocios, manejar grandes volúmenes de información suelen ser un problema que se soluciona con un Sistema Informático, permitiendo así que se minimice el tiempo empleado y los resultados se produzcan con mayor rapidez. Además [14] manifiesta que junto con los datos, los otros dos componentes básicos que constituyen un sistema de información son los usuarios (personal directivo, empleados y en general cualquier agente de la organización empresarial que utilice la información en su puesto de trabajo) y los equipos (informáticos, software, hardware y tecnologías de almacenamiento de la información y de las telecomunicaciones).

1.2.2. SISTEMAS INFORMÁTICOS EN LA WEB

Los Sistemas de Escritorio pueden causar inconvenientes al momento de actualizar el software; además que el tiempo de respuesta del sistema no será el mismo, sino

que dependerá de las características propias de casa computador, impidiendo así la escalabilidad del sistema, aunque en casos particulares convendrá desarrollar un sistema de escritorio, esto dependerá de las funcionalidades que se requieren implementar [17]. Además, según EUATM [18], se denomina páginas Web a documentos que contienen elementos multimedia, además de enlaces de hipertextos. Entrando en la Web a través de cualquier servidor se puede navegar por toda la red, gracias a que la mayoría de las páginas Web contienen enlaces a otras páginas Web que pueden estar localizadas en el mismo servidor o en cualquier otro servidor de Internet.

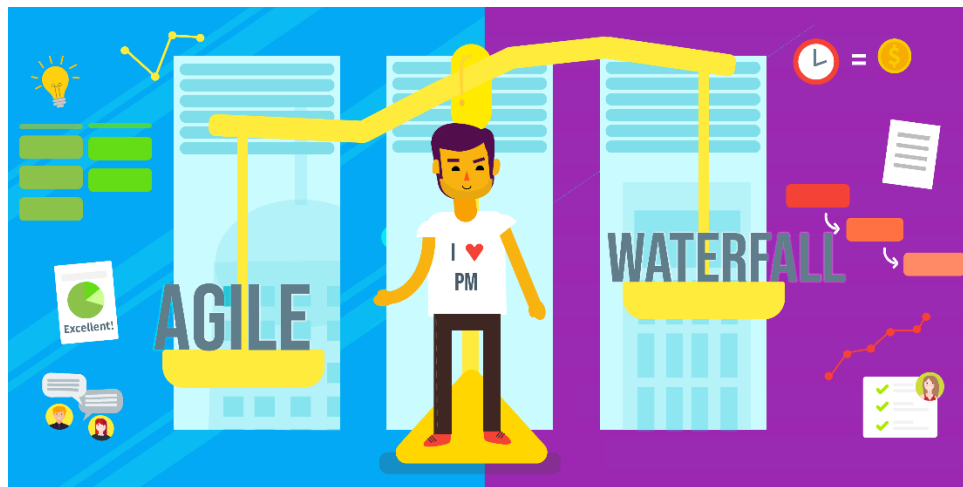
Lo correspondiente a la Web, se basan bajo una estructura Cliente- Servidor, por medio de la cual se atiende a las peticiones de diferentes sitios manejados por los usuarios finales (Información).

1.2.3. METODOLOGÍA ÁGIL EN EL DESARROLLO DE SOFTWARE

Conocemos con los proyectos que no utilizan ningún modelo de desarrollo. Aunque la necesidad de un modelo de desarrollo hace décadas que están firmemente establecidas, es triste comprobar cómo, en nuestro entorno, la desgana y falta de profesionalidad, así como un concepto corto de miras de la gestión empresarial hacen que la mayoría de proyectos de software no apliquen ningún modelo absoluto [19].

En la Figura 1 se observa la balanza entre las metodologías ágiles y metodologías tradicionales, donde claramente se puede evidenciar como las metodologías ágiles están tomando ventaja hacia sus rivales.

Figura 1. Metodología de gestión de proyectos



Fuente: Workep

1.2.4. MODELO ITERATIVO INCREMENTAL

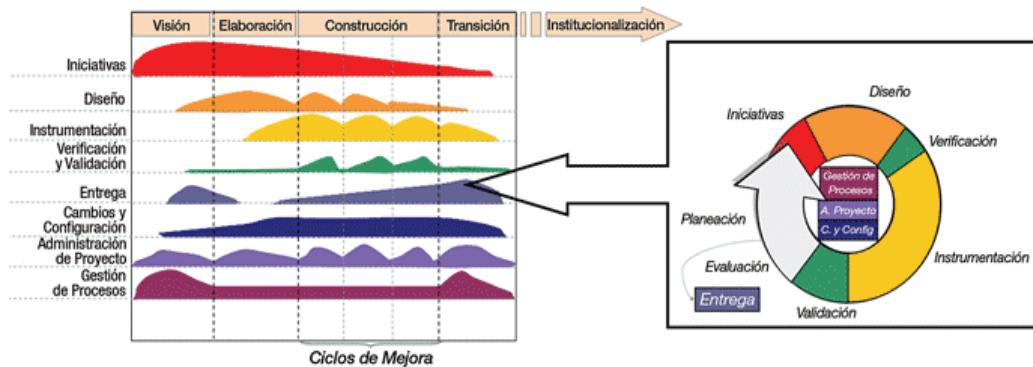
En 2004, [19] dice que los modelos de desarrollo iterativos intentan adaptarse a los cambios, de ahí su idoneidad para el desarrollo de software. Hay varios modelos de desarrollo iterativos. Entre ellos podemos destacar el “Unified Process” y su variante el “Rational Unified Process” que son estándares muy utilizados en la actualidad. Sin embargo, un modelo iterativo que surgió con fuerza es el “Extreme Programming (XP)”. Los modelos iterativos se basan en dividir el proyecto de desarrollo en varias etapas, llamadas iteraciones. Las cuales son cortas (Unas cuantas semanas, excepto en proyectos enormes) y en educación es fija.

Según [20] dicen que El producto se desarrolla por incrementos en el que cada iteración obtiene una versión funcional del producto, de esta forma el sistema se desarrolla poco a poco y obtiene un feedback continuo por parte del usuario.

En 2015, [21] menciona que sobre los modelos iterativos se dice que, en una visión genérica, el proceso se divide en 4 partes: Análisis, Diseño, Código y Prueba. Sin embargo, para la producción del software, se usa el principio de trabajo en cadena o “Pipeline”, utilizado en muchas otras formas de programación. Con esto se mantiene al cliente en constante contacto con los resultados obtenidos en cada incremento. Es el mismo cliente el que incluye o desecha elementos al final de cada incremento a fin de que el software se adapte mejor a sus necesidades reales. El

proceso se repite hasta que se elabore el producto completo como se puede observar en la Figura 2.

Figura 2: Modelo Iterativo Incremental



Fuente: SG Buzz.

En el 2015 [22] clasifica las ventajas y desventajas:

Ventajas:

- Una de las principales ventajas que ofrece este modelo es que no hace falta que los requisitos estén totalmente definidos al inicio del desarrollo, sino que se pueden ir refinando en cada una de las iteraciones.
- Igual que otros modelos similares tienen las ventajas propias de realizar el desarrollo en pequeños ciclos, lo que permite gestionar mejor los riesgos, gestionar mejor las entregas.

Inconvenientes:

- La primera de las ventajas que ofrece este modelo, el no ser necesario tener los requisitos definidos desde el principio, pueden verse también como un inconveniente ya que pueden surgir problemas relacionados con la arquitectura.

1.2.5. CICLO DE VIDA EN EL DESARROLLO DE SOFTWARE.

[23] indica que un modelo de ciclo de vida es un marco de referencia que contiene los procesos, las actividades y las tareas involucradas en el desarrollo, la explotación y el mantenimiento de un proceso de software, abarcando la vida del sistema desde la definición de los requisitos hasta la finalización del mismo. El ciclo de vida abarca toda la vida del sistema, desde su concepción hasta cuando no se utiliza. Además, a veces se habla de ciclo de desarrollo, el cual es un subconjunto

del anterior pues éste empieza en el análisis y finaliza con la entrega del sistema al usuario.

Figura 3: Ciclo de vida del desarrollo de software



Fuente: Software Design Strikers

El ciclo de vida del desarrollo de software se expresa claramente en la Figura 3, donde se observa sus etapas y recorrido hasta completar el producto.

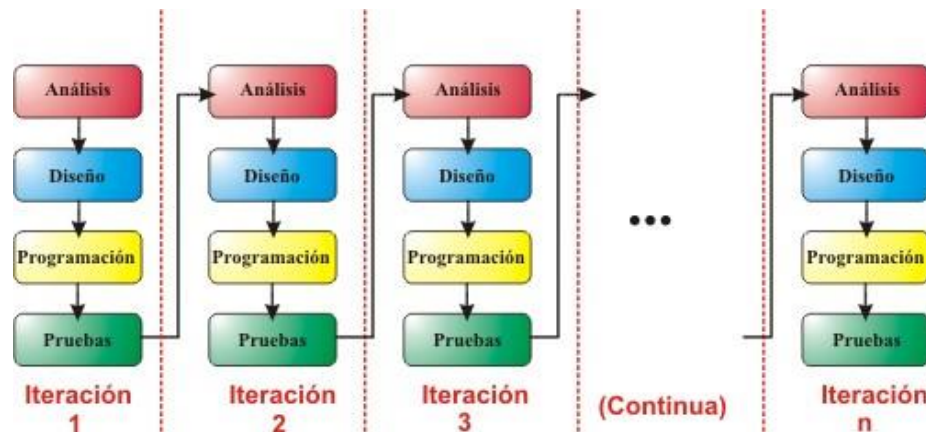
1.2.6. FASES DEL MODELO INCREMENTAL

En el 2017, [24] definen las fases del Modelo Incremental:

- **Requerimientos:** son los objetivos centrales y específicos que persigue el proyecto.
- **Definición de las tareas y las iteraciones:** teniendo en cuenta lo que se busca, el siguiente paso es hacer una lista de tareas y agruparlas en las iteraciones que tendrá el proyecto. Esta agrupación no puede ser aleatoria. Cada una debe perseguir objetivos específicos que la definan como tal.
- **Diseño de los incrementos:** establecidas las iteraciones, es preciso definir cuál será la evolución del producto en cada una de ellas. Cada iteración debe superar a la que le ha presidido. Esto es lo que se denomina incremento.
- **Desarrollo del incremento:** posteriormente se realizan las tareas previstas y se desarrollan los incrementos establecidos en la etapa anterior.

- **Validación del incremento:** al término de cada iteración, los responsables de la gestión del proyecto deben dar por buenos los incrementos que cada una de ellas ha arrojado. Si no son los esperados o sí ha habido algún retroceso, es necesario volver la vista atrás y buscar las causas de ello.
- **Integración de incrementos:** una vez son validados, los incrementos dan forma a lo que se denomina línea incremental o evolución del proyecto en su conjunto. Cada incremento ha contribuido al resultado.
- **Entrega del producto:** cuando el producto en su conjunto ha sido validado y se confirma su correspondencia con los objetivos iniciales, se procede a la entrega final.

Figura 4: Fases del modelo iterativo incremental



Fuente: Pulsodigital

En la Figura 4 se puede apreciar las fases del modelo iterativo incremental, siendo muy utilizada por desarrolladores para programar sus proyectos por su solidez y apoyo durante el ciclo de desarrollo.

1.2.7. LENGUAJE DE PROGRAMACIÓN

Se conoce como lenguaje de programación a un conjunto formal que mediante instrucciones un desarrollador (programador) puede crear funcionalidades a través de acciones consecutivas con datos y algoritmos, buscando controlar el comportamiento físico y lógico de una máquina (computador). De acuerdo a [25] es un idioma artificial diseñado para expresar computaciones que pueden ser llevadas a cabo por máquinas como las computadoras, pueden usarse para crear

programas que controlen el comportamiento físico y lógico de una máquina, para expresar algoritmos con precisión, o como modo de comunicación humana.

1.2.8. PYTHON

En el 2011 [26] dicen que el lenguaje de programación Python se está estableciendo como uno de los lenguajes más populares para la informática científica. Gracias a su naturaleza interactiva de alto nivel y su ecosistema maduro de bibliotecas científicas, es una opción atractiva para el desarrollo algorítmico y el análisis de datos exploratorios. Sin embargo, como un lenguaje de propósito general, se utiliza cada vez más no solo en entornos académicos sino también en la industria.

Según [27] menciona que Python es un modo en que los creadores de programas puedan expresar cómo creen que deben desarrollarse las conversaciones. Durante la programación en Python se comprende como los programadores que lo usan pueden conversar con los usuarios de sus programas.

En el 2016 [28] dice que Python es un lenguaje de programación interpretado e interactivo, capaz de ejecutar en una gran cantidad de plataformas. Se desarrolla como un proyecto de código abierto, administrado por PYTHON Software Foundation.

En la Figura 5 mostrada a continuación se puede observar un ejemplo sencillo realizado en Python:

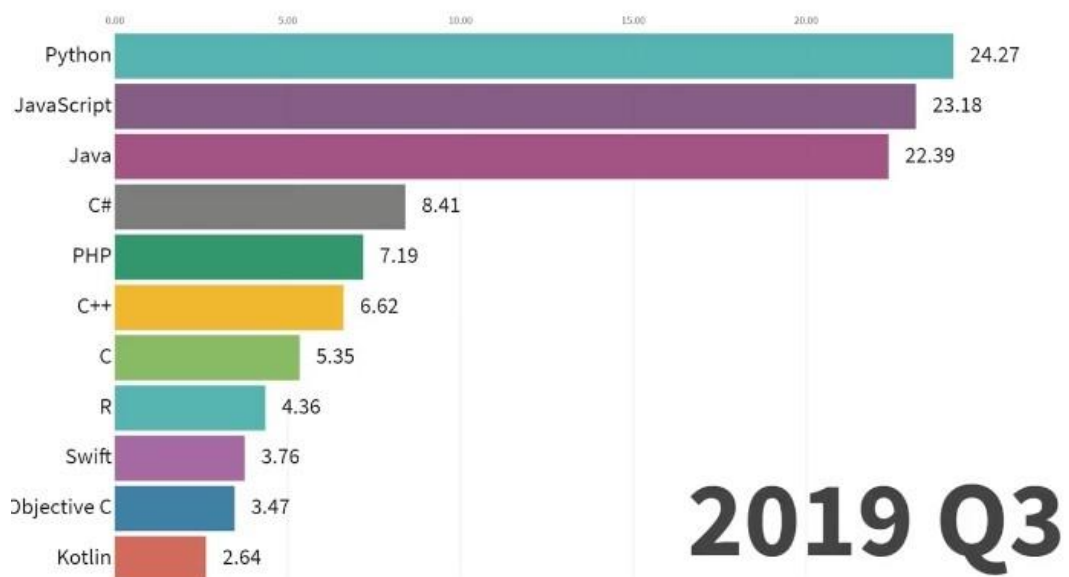
Figura 5: Metodología de Mapeo auto-organizado (SOM)

```
if 1900 < year < 2100 and 1 <= month <= 12 \  
and 1 <= day <= 31 and 0 <= hour < 24 \  
and 0 <= minute < 60 and 0 <= second < 60: # Looks like a valid date  
    return 1
```

Fuente: *The Python Language Reference*

En el ranking de lenguajes de programación más populares para el año 2019 se puede observar en la Figura 6 que se muestra a continuación, como Python se encuentra liderando, esto se debe principalmente por los resultados obtenidos en investigaciones utilizando este lenguaje de programación siendo fiable para el manejo de datos y su procesamiento.

Figura 6: Lenguajes de programación más populares



Fuente: WinDev

1.2.9. FRAMEWORK

Se puede definir a un framework como un armazón, que vendría a ser como una estructura el cual contiene técnicas mediante la utilización de todos los elementos que sean necesarios para beneficio del ser humano. Python es un lenguaje considerado para tomar en cuenta en el desarrollo de aplicaciones cliente, aplicaciones web y sistemas distribuidos, mientras que no es adaptable [28].

1.2.10. FRAMEWORK WEB

Un framework web es una estructura que se encuentra definida que permite la reutilización de sus componentes, los cuales facilitan la creación de estas aplicaciones permitiendo ahorrar tiempo y mantenimiento [28]. En la Figura 7 se encuentra los framework en auge actualmente.

Figura 7: Framework de aplicaciones web



Fuente: ScienceSoft

1.2.11. DESARROLLO DE APLICACIONES WEB MEDIANTE LA UTILIZACIÓN DE FRAMEWORKS

En el desarrollo de software, un framework es una composición conceptual y tecnológica con un soporte bien definido, habitualmente con módulos de software concreto, en base a la cual otro proyecto de software puede ser fácilmente organizado y desarrollo [28].

La arquitectura de los frameworks interactúa bajo el Modelo Vista Controlador (MVC), para que los desarrolladores dividan la documentación.

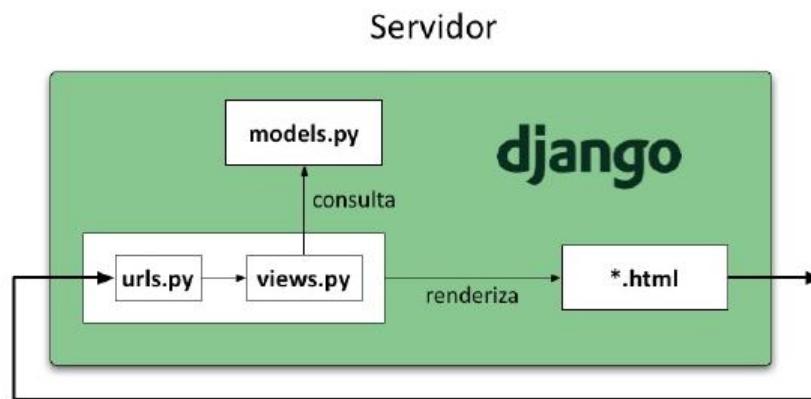
Las ventajas que posee la arquitectura es que permite la separación de sus componentes, es decir la implementación se la realiza de manera separada, al igual que el API está muy bien definida ya que es clara en su uso y se produce su tiempo de ejecución dependiendo del modelo y de sus vistas [28].

- **Modelo:** es aquel que es realizado por el desarrollador y contiene todos los datos, es decir toda la información y la funcionalidad del programa.
- **Vista:** es aquella que permite gestionar como los datos se presentan, es decir la interacción final con la interfaz, la cual debe ser amigable con el usuario o cliente.,

- **Controlador:** Toda la información requerida es enviada al gestor de base de datos para ser guardada, es decir controla el acceso a los datos y de esta manera el contenido es de forma estática y dinámica a la vez.

Lo mencionado referente a la utilización de frameworks en el desarrollo de aplicaciones web se representa gráficamente en la Figura 8, con Django como autor con su estructura interna durante el ciclo de desarrollo.

Figura 8: Django: Esquema interno



- Modelo = Model
- Vista = Template
- Controlador = View

Fuente: Slideshare

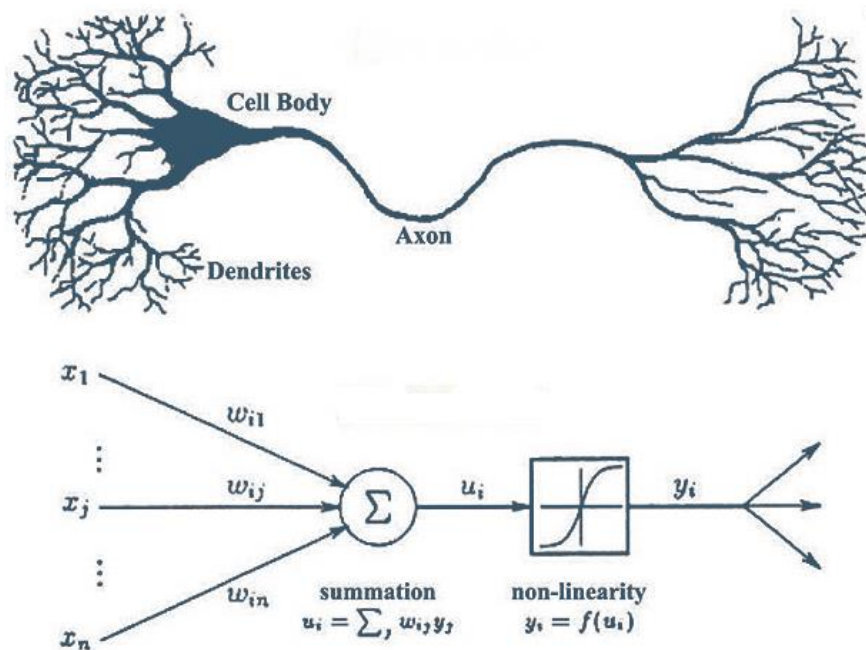
1.2.12. DJANGO

Es un framework web de código abierto escrito en Python que permite construir aplicaciones web más rápido y con menos código, fue inicialmente desarrollado para gestionar aplicaciones web de páginas orientadas a noticias de World Online, más tarde se liberó bajo licencia BSD [28]. Django respeta la modelo vista controlador. Contiene un conjunto de componentes que permite desarrollar sitios web de manera más fácil y rápida. En contexto [29], manifiesta que Django fue inventado para cumplir con estas nuevas ambiciones. Django le permite construir sitios profundos, dinámicos e interesantes en un tiempo extremadamente corto. Django está diseñado para permitirle concentrarse en las partes divertidas e interesantes de su trabajo mientras alivia el dolor de las partes repetitivas.

1.2.13. RED NEURONAL

Es un sistema de procesadores paralelos conectados entre sí en forma de grafos dirigidos. Esquemáticamente cada elemento de procesamiento (neuronas) de la red se presenta como nodo. Estas conexiones establecen una estructura jerárquica que, tratando de emular la filosofía del cerebro, busca nuevos modelos de procesamiento para solucionar problemas concretos del mundo real. Una definición simplificada sobre los mapas topológicos podría ser que, en una correspondencia que respete la topología, las unidades que se encuentran físicamente próximas entre sí van a responder a clases de vectores de entrada que, análogamente, se encuentren cerca unos de otros. Los vectores de entrada de muchas dimensiones son representados sobre el mapa bidimensional, de tal manera que se mantenga el orden natural de los vectores de entrada [30].

Figura 9: Redes neuronales o el arte de imitar el cerebro



Fuente: Magiquo

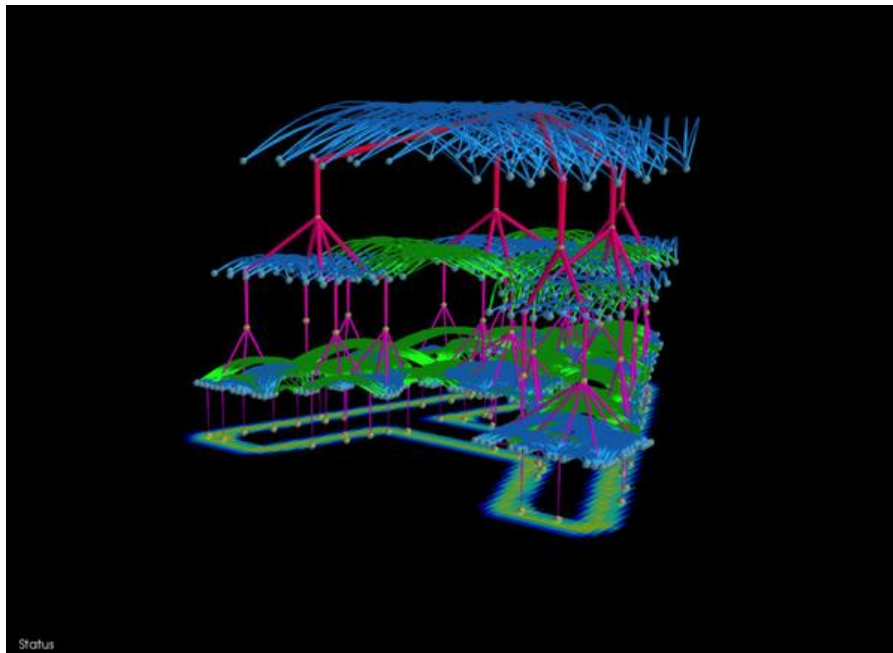
La Figura 9 representa una comparación grafica entre una red neuronal cerebral con una red neuronal artificial creada para el estudio y análisis de datos, en donde hace mención que una red neuronal se refiere al arte de imitar el cerebro.

1.2.14. RED NEURONAL ARTIFICIAL

Las Redes Neuronales Artificiales (RNA) o sistemas conexionistas son sistemas de procesamiento de la información cuya estructura y funcionamiento están inspirados en las redes neuronales biológicas. Consisten en un conjunto de elementos simples de procesamiento llamados nodos o neuronas conectadas entre sí por conexiones que tienen un valor numérico modificable llamado peso [31].

La actividad que una unidad de procesamiento o neurona artificial realiza en un sistema de este tipo es simple. Normalmente, consiste en sumar los valores de las entradas (inputs) que recibe de otras unidades conectadas a ella, comparar esta cantidad con el valor umbral y, si lo iguala o supera, enviar activación o salida (output) a las unidades a las que esté conectada. Tanto las entradas que la unidad recibe como las salidas que envía dependen a su vez del peso o fuerza de las conexiones por las cuales se realizan dichas operaciones [31].

Figura 10: Red neuronal armada por científicos



Fuente: Sophimania

Se puede observar un ejemplo claro de una red neuronal armada por científicos para el trato de datos y su interpretación en la Figura 10.

1.2.15. MAPAS AUTO-ORGANIZADOS

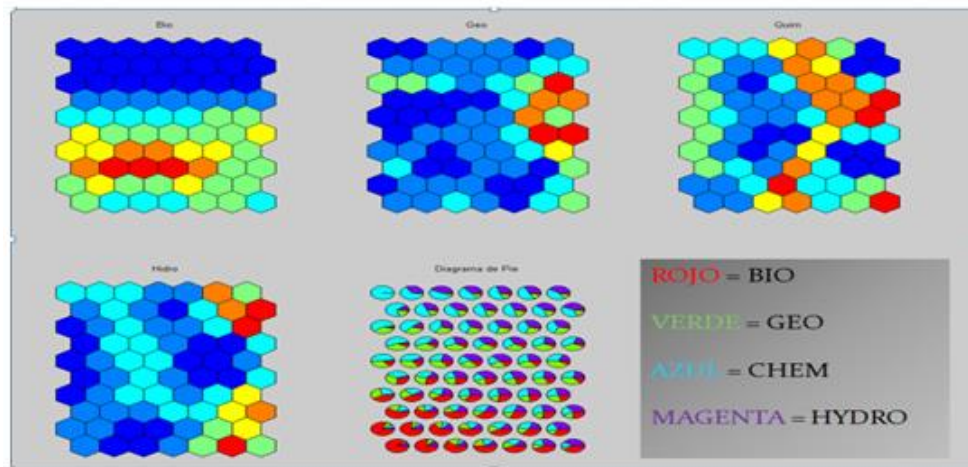
En 1982, [3] presentó un modelo de red denominado mapas auto-organizados o SOM (Self-Organizing Maps), basado en ciertas evidencias descubiertas a nivel cerebral. Este tipo de red posee un aprendizaje no supervisado competitivo. No existe ningún maestro externo que indique si la red neuronal está operando correcta o incorrectamente porque no se dispone de ninguna salida objetivo hacia la cual la red neuronal deba tender.

Con el algoritmo de los Mapas Auto-Organizados (SOM o Self-Organizing Maps), la información de entrada se organiza automáticamente, lo que permite visualizar relaciones importantes entre los datos, a través de mapas bidimensionales de conceptos. Este aspecto es relevante dentro de las disciplinas relacionadas con el descubrimiento de información en grandes bases de datos (Knowledge Discovery in Databases, KDD). Al respecto, es importante mencionar el trabajo de Polanco y colaboradores, del Institut d'Information Scientifique et Technique (INIST) de Francia. Este equipo de trabajo ha desarrollado un sistema llamado Neurodoc, que se basa en el algoritmo SOM de las redes neuronales artificiales, con resultados interesantes [30].

Los Mapas Auto-organizados han sido de utilidad para descubrir y profundizar en el análisis de datos. Se han empleado en el análisis bibliométrico de grandes volúmenes de información, contenida en bases de datos Sotolongo, Guzmán, Carrillo (2001), además, se han empleado en estudios de la ciencia de la salud humana para el mapeo de informaciones cerebrales [32].

La Figura 11 representa la visualización de información clasificada, agrupada y memorizada por un algoritmo que permite interpretar la data de forma gráfica, en este ejemplo se observa los resultados obtenidos de la vigilancia tecnología de ciclos bioquímicos.

Figura 11: Vigilancia Tecnológica de Ciclos Bioquímicos



Fuente: *Journal of technology management & innovation*

1.2.16. MINERÍA DE DATOS

La minería de datos puede definirse inicialmente como un proceso de descubrimiento de nuevas y significativas relaciones, patrones y tendencias al examinar grandes cantidades de datos. La disponibilidad de grandes volúmenes de información y el uso generalizado de herramientas informáticas ha transformado el análisis de datos en minería de datos o Data Mining como se los conoce también esta temática [33].

1.2.17. MINERÍA DE TEXTO

Se refiere al examen de una colección de documentos y el descubrimiento de información no contenida en ningún documento individual de la colección; en otras palabras, trata de obtener información sin haber partido de algo. Debido a que el 80 por ciento de la información de una compañía está almacenada en forma de documentos, las técnicas como la categorización de texto, el procesamiento de lenguaje natural, la extracción y recuperación de la información o el aprendizaje automático, entre otras, apoyan al Text Mining (minería de texto) [34].

1.2.18. PREDICCIÓN

De acuerdo a [35] la predicción es plenamente útil cuando se enmarca dentro del proceso de ayuda a la toma de decisiones, es así que algunos trabajos basados en modelos dinámicos autorregresivos de series temporales y modelos de paseo

aleatorio han demostrado empíricamente que la predicción de resultados basada en el retarde de un periodo de la variable es un buen predictor de ejercicios futuros.

1.2.19. METODOLOGÍA

Se conoce como metodología al conjunto de técnicas empleadas para obtener resultados a partir de un requisito, en este contexto existen metodologías científicas para investigaciones, metodologías para el desarrollo de software, como en nuestro caso metodologías para el diseño e implementación de una necesidad, entre otras muchas. De acuerdo a [36] menciona que el concepto de metodología es propio de la ciencia, sin embargo, se suele aplicar en contextos no científicos, además en la práctica una metodología con carácter científico se pone en funcionamiento en distintas etapas.

1.2.20. LIBRERÍA SOMPY

Es una biblioteca de Python desarrollada para el mapeo autoorganizado (SOM), cuya estructura es similar a la somtoolbox de Matlab. Las funcionalidades que tiene son varias, partiendo por el entrenamiento por lotes, que es más rápido que el entrenamiento en línea, aunque depende del tamaño de la data y del tamaño de la cuadrícula en la cual se reflejara los resultados, Sus resultados se los puede evidenciar mediante mapas de calor y dispone de diferentes métodos para la aproximación de funciones y predicciones (principalmente utilizando Sklearn). Esta librería depende de varias librerías adicionales como la numpy, scipy, scikit-learn, numexpr, matplotlib, pandas e ipdb. La instalación se la puede realizar ejecutando el comando `python setup.py install` en la consola [37].

Durante los últimos dos años, la mayoría de las veces, he estado jugando alrededor del concepto de Mapa Organizador Automático (SOM), que se alinea muy bien con nuestras líneas teóricas de argumentación en nuestro grupo de investigación. Sin embargo, en estos días escuchas en todas partes acerca de Python como un lenguaje de programación de alto nivel para la informática científica, mientras que puedes acelerar tu código a través de diferentes métodos como el uso de Cython, la computación paralela en la nube, el procesamiento de múltiples núcleos y muchos cálculos de matriz interesantes métodos por numpy [38].

1.2.21. AUTOAPRENDIZAJE

Las Tecnologías de la Información y las Comunicaciones (TIC a partir de ahora) favorecen, claramente, la idea de que lo verdaderamente importante es la consecución de unos objetivos y un grado óptimo de calidad más que la presencia física en un lugar y tiempo concretos a la vez que permite generar espacios virtuales compartidos (de relación, de investigación de trabajo) [39].

Este proceso hace referencia a aprender por su propia cuenta a partir de eventos suscitados con anterioridad, marcando el punto de partida para nuevos conocimientos. En el campo de informática el autoaprendizaje tiene como ejemplificación la capacidad que tiene la máquina (computador) de recibir información y utilizarlo en eventos futuros dando soluciones efectivas y rápidas a problemáticas en relación.

1.2.22. TOMA DE DECISIONES

La toma de decisiones es fundamental en cualquier actividad humana. En este sentido, todos somos tomadores de decisiones. Sin embargo, tomar una decisión acertada empieza con un proceso de razonamiento constante y focalizado, que puede incluir varias disciplinas como la filosofía del conocimiento, la ciencia, la lógica, y por, sobre todo, la creatividad. Un gerente debe tomar muchas decisiones todos los días. Algunas de ellas son decisiones de rutina mientras otras tienen una repercusión drástica en las operaciones de la empresa donde trabaja [40].

En cuanto a la toma de decisiones en lo que a las TIC corresponde, se tiene mucho que ver con la data existente y como se procesa. Es importante conocer que para una mejor toma de decisiones es necesario tener la data correctamente clasificada y estructurada, es por eso que se utilizan algoritmos para el aprendizaje y muestra de resultados de manera interactiva, permitiendo al usuario comprender claramente la respuesta proporcionada por el sistema que se maneje.

1.2.23. MANEJO DE ENTORNOS VIRTUALES

Según [41] conjunto de herramientas que permiten crear entornos aislados, donde cada uno de ellos es responsable de sus propios paquetes. Es decir, nos permiten crear una sandbox en la que instalar diferentes módulos que serán independientes,

tanto de los que existen globalmente en el sistema (dentro del directorio site-packages), como de otros entornos virtuales. Bastante interesante cierto, y espera a ver lo simple que es su manejo.

[42] dice que, para el manejo de entornos virtuales en Python, necesitamos instalar un módulo llamado virtualenv, el cual es accesible a través de Pip. Ello implica que podremos instalarlo con un gestor de paquetes como easy_install y pip. Una vez instalado tendremos acceso a un script Python con el mismo nombre, el cual puede ser invocado directamente. Esto se crea para el mejor manejo de proyectos y versiones en cuanto a sus paquetes y demás disminuyendo la probabilidad de generar errores durante el desarrollo del proyecto en concreto.

1.2.24. ÁREAS DEL CONOCIMIENTO

Las áreas del conocimiento en educación se refieren a la clasificación de conocimiento científico por áreas para su estudio, este tema se evidencia de forma más clara en la educación superior donde se observa la sistematización por lineamientos para separar las áreas de estudio por carreras y materias.

1.2.25. BASE DE DATOS

Una base de datos es un conjunto de datos almacenados en memoria externa que están organizados mediante una estructura de datos. Una base de datos se puede percibir como un gran almacén de datos que se define y se crea una sola vez, y que se utiliza al mismo tiempo por distintos usuarios. Según [43], antes de existir las bases de datos, los programas debían manejar los datos que se encontraban almacenados en ficheros desconectados y con información redundante. En una base de datos todos los datos se integran con una mínima cantidad de duplicidad. De este modo, la base de datos no pertenece a un solo departamento, sino que se comparte por toda la organización. Además, la base de datos no sólo contiene los datos de la organización, también almacena una descripción de dichos datos. Esta descripción es lo que se denomina metadatos, se almacena en el diccionario de datos o catálogo y es lo que permite que exista lo que se denomina independencia de datos lógica-física, de la que se hablará más adelante.

La esquematización gráfica de una base de datos y su interacción con los sistemas en computadores o dispositivos inteligentes se puede justificar en la Figura 12 mostrada a continuación.

Figura 12: Esquema de Base de Datos



Fuente: Astera Software

1.2.26. POSTGRESQL

Existen tres aplicaciones básicas de productividad de oficina: procesadores de texto, hojas de cálculo y bases de datos. Los procesadores de texto producen documentos de texto críticos para cualquier negocio. Las hojas de cálculo se usan para cálculos y cálculos financieros. Las bases de datos se utilizan principalmente para el almacenamiento y la recuperación de datos. Puede usar un procesador de textos o una hoja de cálculo para almacenar pequeñas cantidades de datos. Sin embargo, con grandes volúmenes de datos o datos que deben recuperarse y actualizarse con frecuencia, las bases de datos son la mejor opción. Las bases de datos permiten el almacenamiento ordenado de datos, la recuperación rápida de datos y el análisis complejo de datos [44].

1.3. FUNDAMENTACIÓN DEL ESTADO DEL ARTE

Se han realizado numerosas investigaciones a nivel mundial en cuanto a clasificación no supervisada de datos refiere, en las mismas llama la atención el mapeo auto-organizado (SOM), debido, que a pesar de estar ya varias décadas en

auge no ha sido considerado adecuadamente hasta la última década por la continua actualización de la tecnología acorde a la era digital lo que conlleva al manejo de grandes cantidades de datos “Big Data”. Es a partir de lo mencionado que resulta útil la implementación de SOM en la visualización de información ya que permite al usuario interactuar adecuadamente con la tecnología. A continuación, se mencionan algunas de las investigaciones más relevantes referentes a IA, RNA, SOM y clasificación de información.

En Alemania [45] propone en su tesis para la obtención de su Maestría que las grandes colecciones de archivos de audio nunca estuvieron tan disponibles, ya que son fácilmente accesibles en línea como en medios de almacenamiento. Al mismo tiempo, las herramientas para navegar, buscar y organizar estas colecciones de archivos de audio son cada vez más inmanejables, es decir no han seguido el ritmo; por lo que en esta tesis propone implementar el Mapa Auto-organizado (SOM), un algoritmo de aprendizaje automático establecido para la reducción de la dimensionalidad y la visualización de los datos, y aplicarlo a la organización de los sonidos.

En Turquía de acuerdo a [46] en su investigación referente al modelado y simulación del uso de la tierra / cambio de cubierta mediante red neuronal artificial de los datos de sensación remota habla acerca del aumento de la población, movilidad y necesidades de los seres humanos afecta significativamente en la dinámica del uso y la cobertura del suelo. En dicho estudio da énfasis a detectar los cambios temporales y espaciales de la cubierta terrestre y la población humana, en el noroeste de Turquía, a través del uso de imágenes satelitales de la cubierta terrestre en los años 1997, 2007 y 2017 para estimar el año 2027 mediante el enfoque RNA. En la cual el resultado del aprendizaje fue del 80.6 % y el valor de validación de corrección fue de 90.1%.

En Estados Unidos se puede apreciar la tesis para la obtención de Master de Ciencias en Informática, realizada por [1], en la cual dice que los SOM han sido bien investigados desde la década de 1980 y se han implementado en C, Fortran, R y Python, siendo Python un lenguaje de alto nivel en el campo del aprendizaje automático sus librerías solo realizan la construcción y visualización de modelos,

es por esto que se tradujo la librería POPSOM originalmente escrito en R la cual es capaz de realizar funcionalidades más allá de las mencionadas, como evaluar la calidad del modelo con métodos estadísticos y trazar distribuciones de probabilidad marginal de las neuronas.

En Cuba por su parte [7] realizaron un estudio comparando las técnicas de IA para el diagnóstico de enfermedades en la ganadería, en la cual mencionan que en la actualidad la medicina (animal y humana) utiliza considerables adelantos que involucran el uso intensivo de alta tecnología como el diagnóstico por imágenes, la robótica, entre otros, especialmente la aplicación de IA las cuales son factibles cuando aprovechan los datos disponibles y la experiencia clínica. El presente trabajo tiene como objetivo, comparar las técnicas de IA para seleccionar la que mejor se ajusta al diagnóstico de enfermedades en la ganadería, cuando se tienen los datos almacenados sobre el comportamiento de las enfermedades que con frecuencia afectan la masa ganadera.

En Brasil [3] proponen la investigación referente a Redes Neuronales Artificiales aplicadas en Biometría Forestal y Modelado, en la cual mencionan la IA a pesar de centrarse en una amplia gama de métricas exitosas, el enfoque de la el enfoque de la Red Neural Artificial (ANN), una técnica similar al sistema nervioso central, ha ganado notoriedad y relevancia con respecto a la clasificación de estándares, estimaciones de parámetros intrínsecos, detección remota, minería de datos y otras posibilidades, este artículo tiene como objetivo realizar una revisión sistemática, que involucra algunos aspectos bibliométricos, para detectar la aplicación de ANN en el campo de la Ingeniería Forestal, particularmente en el pronóstico de los parámetros esenciales para el inventario forestal, analizando la construcción de los ámbitos, la implementación de redes (tipo - clasificación).

En Brasil también se encuentra la investigación realizada por [47] referente a IA, la cual dice que se han desarrollado sistemas de diagnóstico asistidos por computadora con el objetivo de complementar las imágenes de diagnóstico y ayudar al proceso de toma de decisiones terapéuticas. Con el advenimiento de la inteligencia artificial, el "big data" y el aprendizaje automático, estamos avanzando hacia la rápida expansión del uso de estas herramientas en la vida diaria de los médicos, haciendo

que cada paciente sea único, así como liderando la radiología hacia el concepto de multidisciplinario enfoque y medicina de precisión.

En Perú se realizó la investigación acerca de clasificación de datos basada en comprensión por [48], de acuerdo a esta se dice que el incremento del volumen de datos en esta era digital es enorme, la tarea de analizarlos, procesarlos, identificarlos para luego poder clasificarlos y así tener un buen sistema de minería de datos donde poder indexar la información que contiene sin importar la cantidad y el tipo de datos resulta una tarea nada fácil. Debido a esto cada vez se hace más necesario el desarrollo de métodos más efectivos que faciliten esta tarea de forma automática.

Por su parte en Ecuador se encuentra la investigación realizada en la Universidad San Francisco de Quito por [49] en la cual plantean examinar el estado de actividad del volcán Tungurahua (Ecuador) durante el año 2014, mediante el análisis de diferentes eventos sísmicos registrados en una estación geofísica permanente del Instituto Geofísico de la Escuela Politécnica Nacional ubicada en dicho volcán. Por eso en este trabajo se propone agrupar las señales sísmicas de forma no supervisada en búsqueda de posibles nuevas clases de eventos que podrían indicar el inicio de fases eruptivas o que permitan describir mejor el estado del volcán. El análisis como tal se realizó utilizando algunas técnicas de clasificación no supervisada de objetos (k-medias, análisis de arquetipos y mapas auto-organizados) sobre señales sísmicas "discretas", es decir que no superaron más de 2 minutos de duración.

En la Universidad Técnica de Cotopaxi existe falencias en la clasificación de documentos científicos producidos por sus docentes investigadores, esto se debe primordialmente por la gran cantidad de datos que son almacenados y procesados, por lo que se encuentra prudente implementar un método capaz de entrenar, clasificar y visualizar dicha información sin la necesidad de supervisión. En la actualidad se evidencia la existencia de diferentes investigaciones a nivel mundial referentes a técnicas de IA para la clasificación no supervisada y la visualización de la misma. Una de las técnicas más llamativa y eficiente es el Mapeo Auto-organizado (SOM), la cual se acopla perfectamente para dar solución a la problemática propuesta durante la investigación.

1.4. CONCLUSIONES CAPÍTULO I

- Al culminar con el análisis bibliográfico del tema planteado se verifica la importancia de las redes neuronales y la inteligencia artificial, además, del impacto que está teniendo hoy en día.
- El mapeo auto-organizado en la actualidad está siendo utilizado para investigaciones referentes a Redes Neuronales e Inteligencia Artificial, siendo de suma importancia para la visualización de información mediante mapas de calor.
- La Metodología planteada está enfocada para el autoaprendizaje y toma de decisiones, de esta manera se obtendrá mejores resultados al procesar y visualizar la información.

CAPÍTULO II

2. PROPUESTA

2.1. DIAGNÓSTICO

En la Universidad Técnica de Cotopaxi desde hace un tiempo se viene desarrollando un sistema web por nombre ECUCIENCIA, el cual pretende organizar la información de docentes investigadores y estudiar la colaboración de artículos, libros y ponencias científicas de los mismos. La automatización de los procesos según la clasificación por áreas del conocimiento en los datos de producción científica se torna compleja por la gran cantidad de información, sumándole a esto el cuantioso número de docentes investigadores que al presente forman parte de la Universidad Técnica de Cotopaxi en sus diferentes Facultades.

Actualmente dicha clasificación se la realiza manualmente, provocando que esta situación se torne compleja y tardía debido al tiempo empleado en los procesos mencionados, lo que resulta un verdadero dolor de cabeza para el personal encargado de dichos métodos, además la ubicación de dicha información es confuso y riesgoso. Ocasionalmente por lo expuesto es necesario la impresión física de la información para el mejor análisis y clasificación de la misma. Esto genera desperdicios tanto de recursos humanos, económicos y ambientales, adicional el riesgo de deficiencias los cuales serán reflejados en el sistema web, originando pérdida de información y credibilidad en el mismo que directa o indirectamente afectan a la Universidad Técnica de Cotopaxi.

La propuesta para dar solución a los inconvenientes planteados es la creación de un módulo que permita clasificar, organizar y visualizar los datos de producción científica de los docentes investigadores en la Universidad Técnica de Cotopaxi por áreas del conocimiento, automatizando los procesos de acuerdo a las líneas y

sublíneas de investigación en la Universidad Técnica de Cotopaxi, CACES y UNESCO según el campo amplio, específico y detallado. Esto será posible implementando Redes Neuronales Artificiales (ANN), logrando visualizar los datos mediante mapeos auto-organizados (SOM) distinguiéndolos por colores, mejorando considerablemente la interacción con el usuario y el personal encargado del sistema.

Como punto adicional se pretende el módulo tenga la capacidad de tomar decisiones, es decir aprender de acuerdo a los datos procesados, dando al sistema el control para clasificar automáticamente según el nivel de impacto del artículo científico, el cual será posible medirlo de acuerdo a la categorización de la Revista a la que fue indexada proporcionando recompensas a los investigadores que posean más artículos científicos en revistas de prestigio. Para la implementación de lo formulado, se utilizará el método de trabajo alineado a SOM que fue desarrollada acorde a las necesidades presentadas en el proceso de la propuesta planteada, ya que permite tomar decisiones especialmente en situaciones complicadas en las que existen varias alternativas, criterios y puntos de vista. Además, se utilizará la librería SOMPY en el lenguaje de programación Python para el SOM.

2.2. METODOS ESPECÍFICOS DE LA INVESTIGACIÓN

Considerando la problemática y antecedentes se decide utilizar una metodología de desarrollo propia debido a la recopilación de opiniones y resultados de investigaciones planteadas por [1], [38], [45] como principales referentes, en las cuales dan énfasis a la clasificación y visualización de información mediante SOM. Los principales inconvenientes del método SOM que se han presentado en las investigaciones propuestas son los siguientes:

- Los SOM no fueron diseñados para optimizar los resultados, aunque en la actualidad los investigadores se encuentran acoplado esta metodología para este fin.
- La competitividad de las neuronas depende fundamentalmente de la data de entrada.
- La visualización de los mapas con la data entrenada y clasificada se realiza directamente en imágenes estáticas dificultando la implementación en la web.

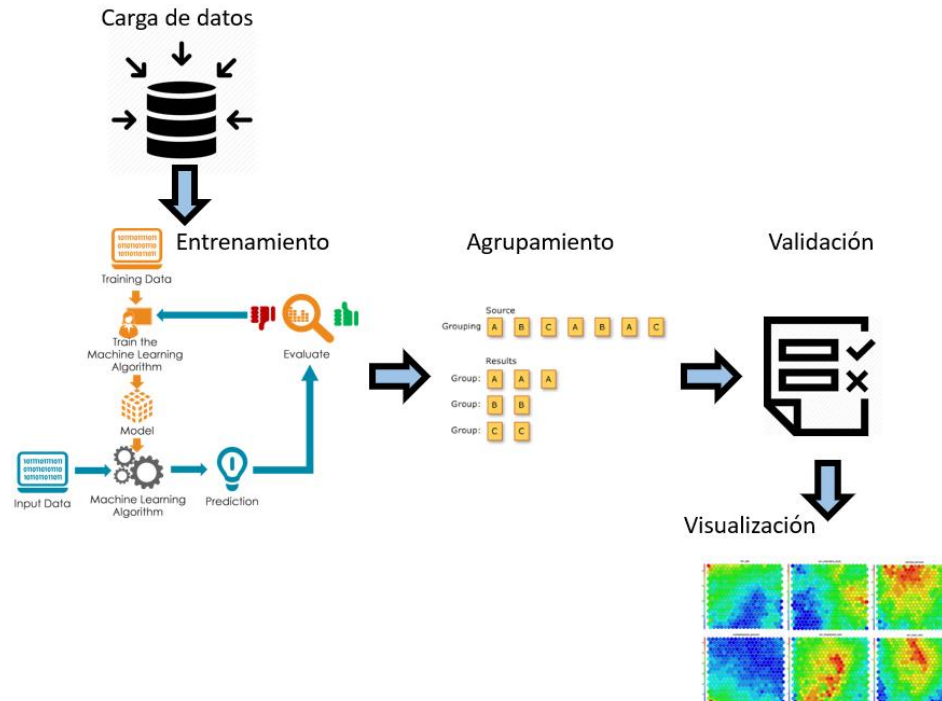
- El tiempo de respuesta entre la carga de datos y visualización es escasamente lenta.
- Los mapas autoorganizados visualizados al culminar con sus respectivos procesos no son completamente estables y tienen ciertas variaciones.

Es por esto que el método SOM propuesto pretende ajustar las falencias encontradas en las indagaciones, además el mismo es diseñado acorde a las necesidades presentadas en esta investigación.

2.2.1. MÉTODO DE TRABAJO ALINEADO A MAPEO AUTO-ORGANIZADO (SOM)

El mapeo auto-organizado (SOM) se refiere a modelos de redes neuronales artificiales basadas en descubrimientos evidenciados a nivel cerebral, por lo que posee un nivel de aprendizaje no supervisados, es decir no necesita un usuario supervisor que verifique el funcionamiento de la red neuronal por lo que se encarga de clasificar, aprender y visualizar los resultados a partir de una data de entrada. El aprendizaje lo realizan las neuronas compitiendo entre ellas para llegar al resultado más óptimo para mostrar al usuario mediante mapas de calor. El método de trabajo está enfocado desde el análisis de datos hasta la visualización del usuario como se muestra en la Figura 13 donde se puede observar el Método de trabajo alineado a SOM planteado para obtener los resultados esperados en la investigación.

Figura 13. Método de trabajo alineado al Mapeo Auto-organizado (SOM).



Elaborado por: Alex Geovanny Rivera Chasiquiza

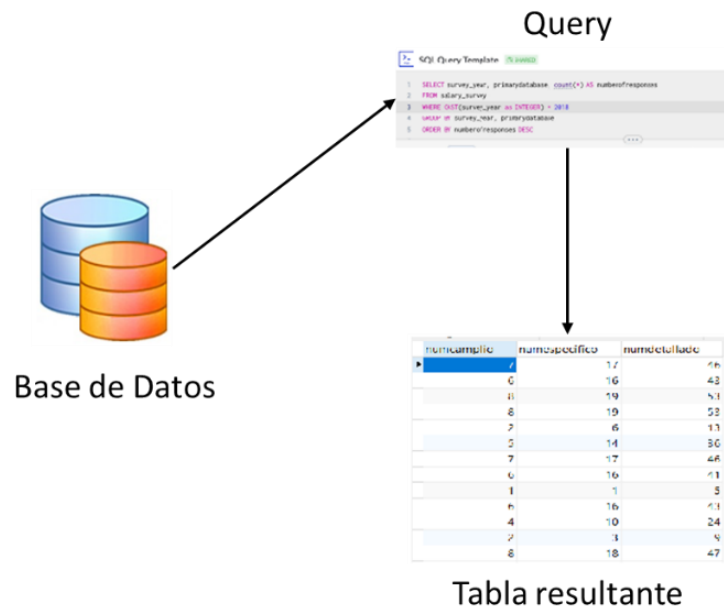
En el aprendizaje competitivo las neuronas compiten unas con otras con el fin de llevar a cabo una tarea dada. Se pretende que cuando se presente a la red un patrón de entrada, sólo una de las neuronas de salida (o un grupo de vecinas) se active.

Un modelo SOM está compuesto por dos capas de neuronas. La capa de entrada (formada por Neuronas, una por cada variable de entrada) se encarga de recibir y transmitirá a la capa de salida la información procedente del exterior. La capa de salida (formada por M neuronas) es la encargada de procesar la información y formar el mapa de rasgos. Normalmente, las neuronas de la capa de salida se organizan en forma de mapa bidimensional. A continuación, se especifican las etapas importantes de dicho método:

ETAPA 1. Carga de datos: Es importante tener la base de datos estructurada y normalizada para el correcto análisis de la misma. Los artículos científicos, libros, ponencias y revistas son sumamente importantes para la carga de datos de acuerdo a las líneas y sublíneas de investigación, además del campo amplio, específico y detallado en cuanto a la UNESCO y CACES se refieren. Estos conceptos

mencionados deben constar en la base de datos del el Sistema ECUCIENCIA, para los siguientes pasos que el método de trabajo plantea hasta llegar a los resultados esperados. En la Figura 14 mostrada a continuación se interpreta esta preparación de datos.

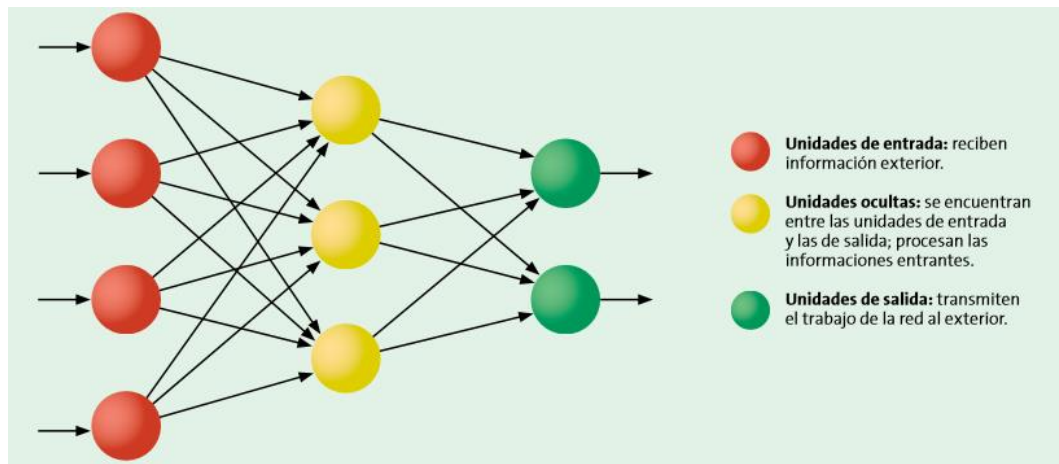
Figura 14. Carga de Datos



Elaborado por: Alex Geovanny Rivera Chasiquiza

ETAPA 2. Entrenamiento: Los datos obtenidos en la clasificación previa son vitales como se lo menciono debido a que son el punto de partida para entrenarlos y puedan ser capaces de tomar decisiones. Con el entrenamiento de los datos mediante redes neuronales artificiales se puede obtener resultados que serán tratados para llegar a la visualización en los mapas de calor. Para ser más específico en este proceso se muestra en la Figura 15 gráficamente como realiza el proceso de entrenamiento y entrega de resultados.

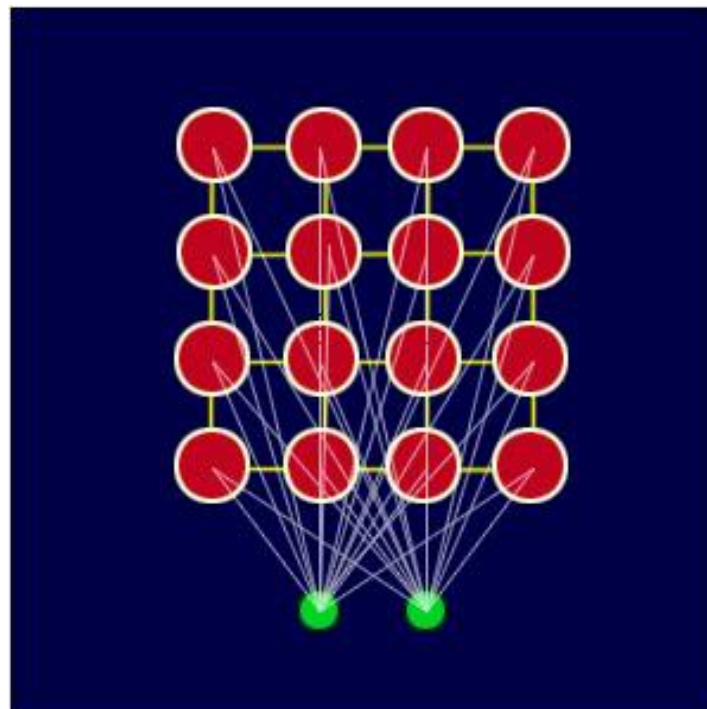
Figura 13. Entrenamiento de las neuronas



Elaborado por: Andrea Vedaldi

ETAPA 3. Agrupamiento: A los datos expuestos del entrenamiento en las redes neuronales se los puede agrupar para tener un resultado más óptimo y veraz. Este proceso al igual que los anteriores son importantes para poder continuar con el debido entrenamiento y de esta manera permitir los resultados sean más efectivos. En la Figura 16 mostrada a continuación se puede apreciar de mejor manera en lo que consiste el agrupamiento.

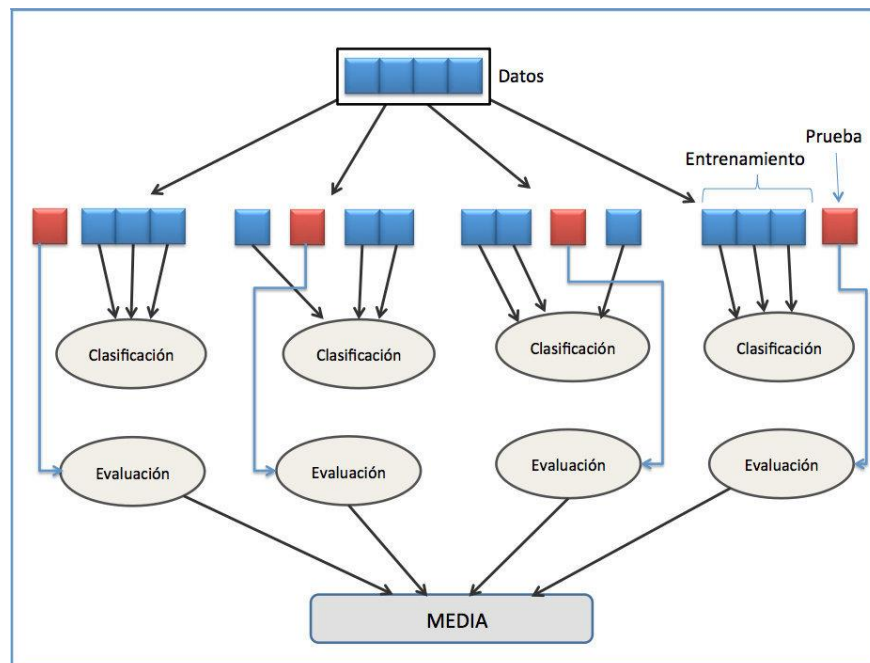
Figura 14. Agrupamiento de los datos resultantes



Elaborado por: Fernando Sancho Caparrini

ETAPA 4. Validación: El siguiente aspecto a tomar en cuenta es la validación de los resultados, este criterio se lo debe realizar acorde a parámetros y algoritmos que verifiquen los resultados, permitiendo asegurar que estos son fiables y lo suficientemente confiables para mostrar al usuario receptor de la data. Un método valido para redes neuronales artificiales es la validación cruzada, la cual evalúa los resultados estadísticos. Cabe recalcar que la librería Sompy desarrollada en Python se encarga de esta validación de los resultados antes de devolver la información que será procesada para su visualización. En la Figura 17 se presenta el gráfico correspondiente a validación cruzada.

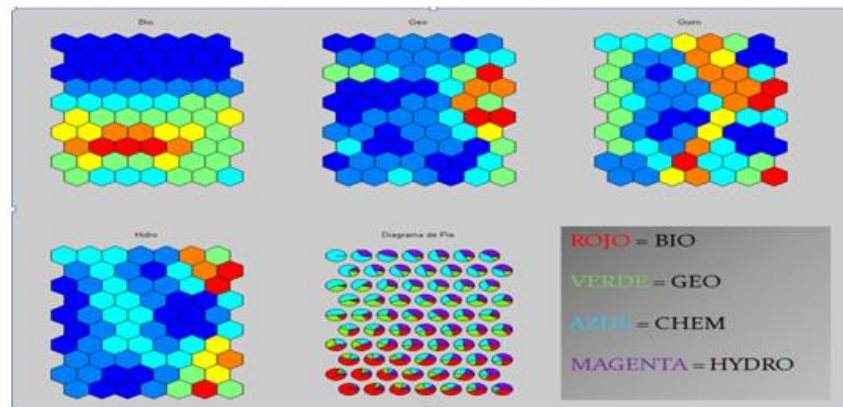
Figura 15. Validación de la información



Fuente: Validación cruzada De Joan.domenech91

ETAPA 5. Visualización: Por último, nos queda mostrar los resultados, esta fase es prudente realizarla mediante mapas de calor ya que permiten exponer con claridad las derivaciones obtenidas, haciendo que el sistema sea interactivo y funcional para su uso. Por lo que el usuario final quedara satisfecho con lo obtenido como se muestra en la Figura 18 expuesta a continuación.

Figura 16. Visualización de información procesada



Elaborado por: Carlos Felipe Gaitán Ospina

Los modelos que representan entrenamiento de neuronas, son el resultado de investigaciones arduas relacionadas al manejo de grandes cantidades de datos, los cuales son clasificados, analizados y entrenados para poder dar resultados visibles. Este tema se encuentra en pleno desarrollo para la creación de software en particular ya que se evidencia un gran cumulo de resultados fiables que dan el punto de partida para que demás investigadores se inmiscuyan y puedan dar un valor agregado a lo que presentan a sus usuarios, además que la data al ser tratada es capaz de auto aprender y tomar decisiones no asistidas provocando que los sistemas sean aún más interactivos e íntegros.

En la actualidad como se menciona el manejo de datos es algo imprescindible en la vida cotidiana y aún más en el desarrollo de software, por lo que exige que las grandes cantidades de datos sean analizadas. En el caso de los datos de producción científica en la Universidad Técnica de Cotopaxi, es importante se ejecute un análisis complejo de los artículos científicos, ponencias, libros y revistas, ya que por este medio se podrá obtener una clasificación seria que ratificará los resultados como verídicos. Por este motivo es conveniente la utilización de este método de trabajo alineado a SOM en el proceso de dar solución a la problemática planteada en la investigación.

- **Propiedades de los sistemas autoorganizados**

Las redes neuronales artificiales con aprendizaje no supervisado son puestas en práctica a menudo, obteniendo como resultados el reconocimiento de patrones que

generan internamente a partir de los datos de entrada las neuronas compitan. Este proceso es el que se pretende plasmar con los datos de producción científica en la Universidad Técnica de Cotopaxi. Para que las neuronas se activen debe existir una relación entre los individuos y artículos, libros y ponencias; para mediante este resultado se estructure un patrón interrelacionado entre sí en la carga de datos.

Las conexiones de la red neuronal artificial son importantes en el procesamiento de patrones, debido a que en el caso de no existir una conexión las neuronas quedarán indispuestas en el aprendizaje, reflejando resultados menos probables y causando problemas en la visualización para el usuario. Cabe mencionar que las poblaciones con mayor densidad en la red neuronal artificial serán quienes más perspectivas en el aprendizaje aportarán, es decir serán las que de acuerdo a los artículos, ponencias, libros y revistas aportarán resultados para la visualización.

El método de trabajo alineado a SOM pretende mejorar el proceso de aprendizaje de datos para una mejor visualización de resultados, tomando como prioridad los datos de ingreso, es por esto importante la correcta estructura y normalización de la base de datos. Una vez verificado la data de carga las redes neuronales artificiales encontrarán menos trabas para el proceso de aprendizaje, agrupamiento y validación de los mismos; esto generará un valor extra que se expresa en la visualización. Es importante tener en cuenta las recomendaciones propuestas en este capítulo para el manejo de datos. En el caso de las redes neuronales artificiales referentes a los datos de producción científicos en la Universidad Técnica de Cotopaxi en el sistema ECUCIENCIA se tomará las relaciones entre las tablas mencionadas que pertenecen a la base de datos y la clasificación se realizará a través de vistas, con esto el sistema se encontrará en constante aprendizaje proporcionando resultados siempre reales y fiables.

2.2.2. METODOLOGÍA DE DESARROLLO

Durante la investigación se plantea el método de trabajo SOM (Mapeo Autoorganizado) como una metodología más precisa para el proceso de toma de decisiones y machine learning, partiendo desde la carga de datos hasta culminar con los resultados de las etapas planteadas en su visualización mediante mapas de calor. Al ser la metodología de desarrollo de software sumamente importante en el

proceso de programación de aplicaciones es prudente que dicha metodología sea realizada por cada Etapa del Método SOM propuesto en la Figura 13. A continuación se explica detalladamente las Fases de la metodología de desarrollo de software para posteriormente ponerlas en práctica.

- **Etapas de la metodología de desarrollo de software**

El desarrollo de software se ha convertido en uno de los campos más competitivos en la actualidad, debido a esto las metodologías de desarrollo se encuentran en constante evolución teniendo como objetivo mejorar y optimizar el producto resultante para una mejor interacción con el usuario. Debido a que el ciclo de vida del software es variable es importante identificar la metodología idónea. De acuerdo a [50] los modelos ágiles hacen hincapié que la estimación de estados de usuarios y sus respectivas tareas deben realizarlas por el equipo de desarrollo. También, manifiesta que el modelo ágil establece una filosofía para guiar el sistema de modelado de software, además presenta algunos principios en la siguiente tabla.

Tabla 2: Principios del Modelado Ágil

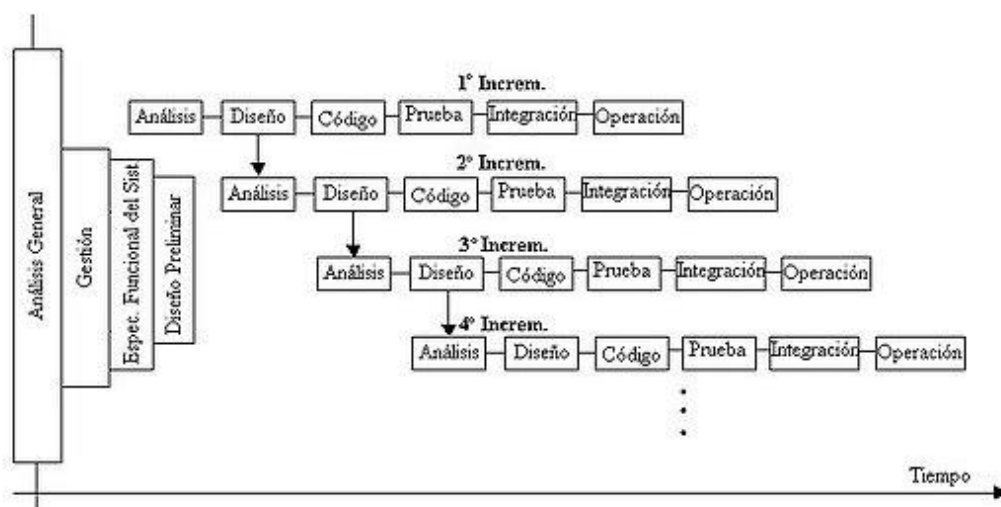
Modelar con un propósito	Solo crear diagramas, documentación y especificaciones si hay un objetivo específico para eso.
El uso de varios modelos	Elegir algunas de las múltiples formas de modelado existentes que sean necesarias y representativas.
Mantenga solo lo necesario	Actualizar los modelos a lo largo de todo el proyecto conforme los cambios en los requisitos generan un trabajo considerable, por lo que sólo lo que es realmente relevante debe mantenerse.
El contenido es más importante que la representación	La preocupación debe ser la de transmitir la idea y no el formalismo de un modelo, es decir, la información contenida en los artefactos debe ser suficiente para la situación actual.

<p>Conozca a los modelos y herramientas</p>	<p>Saber utilizar adecuadamente las plantilla y herramientas de ayuda en la decisión de la forma de modelar un problema. Los diagramas tienen muchas características y limitaciones que se desechan según sea necesario.</p>
<p>Adaptar localmente</p>	<p>El modelo debe ser adoptado a las necesidades del proyecto y el equipo.</p>

Fuente: Ambler

Según lo mencionado en la tabla 1, para el desarrollo del módulo de visualización, clasificación y automatización en datos de producción científica de docentes investigadores de la Universidad Técnica de Cotopaxi mediante mapeo auto-organizado se utilizará la metodología de desarrollo iterativo e incremental, ya que el proyecto se proyecta en bloques temporales, es decir planificar pasos lógicos en cada iteración hasta llegar a una iteración final siguiendo los cronogramas establecidos en el periodo de desarrollo. Este modelo está basado en una secuencia de modelos en cascada aplicados hasta llegar al producto (software) final. En la figura 19 se muestra un esquema del modelo iterativo incremental con sus diferentes etapas.

Figura 17. Modelo iterativo incremental.



Como se observa en la Figura 19 se puede apreciar claramente esquema del modelo iterativo incremental y sus etapas. Al culminar cada etapa comienza una nueva iteración y así proseguir con el cronograma planificado, esto se realiza hasta llegar a la solución

- **Análisis:** Requerimientos necesarios para el desarrollo del software.
- **Diseño:** Bosquejo del software.
- **Código:** Desarrollo y codificación de la solución:
- **Prueba:** Testeo de la solución realizada.
- **Integración:** Sacar a producción la solución.
- **Operación:** Mantenimiento de la solución presentada.

2.2.3. ESTRATEGIAS DE IMPLEMENTACIÓN, COMPARATIVAS Y ADOPCIÓN TECNOLÓGICA:

- **Estrategias de implementación:**

Al culminar con una solución, es común encontrarse con errores al implementar la solución en el cliente, siendo estos tanto de software, hardware y en ciertas ocasiones humanas. De acuerdo a [51] afirma que entre el 55% y el 75 % de todos los proyectos ERP fallan en su implementación. En un panorama como este el bienestar financiero y cultural de todo negocio se encuentra en juego y los costos asociados a las fallas pueden ser catastróficos. Es por esta razón implementar estrategias para tener éxito, tanto desde el desarrollo del sistema hasta su implementación. Es así que [51] sugiere aplicar las siguientes 7 estrategias para triunfar en la implementación:

1. Haz que todos se involucren: Es necesario que todos los involucrados comprendan la necesidad del cambio y el valor que el sistema aportará.
2. Establece los objetivos del proyecto: Es necesario que los objetivos estén claros y alcanzables.
3. Prepárate para ser flexible: Identificar los cambios clave en los sistemas, procesos y estructura organizacional.
4. La comunicación es clave: Asegúrate de contar con canales de comunicación multidireccionales que garanticen la información.

5. Presta atención a las modificaciones / personalizaciones: Tener en cuenta las personalizaciones y las expectativas claras respecto a estas.
6. Incluye un presupuesto realista: Tener en cuenta todos los costos asociados al software y su implementación.
7. Elige un partner con experiencia: Aumenta las posibilidades de éxito.

Para la correcta implementación del sistema se debe tomar en cuenta las estrategias mencionadas, de esta manera se minimiza los problemas tanto económicos y de tiempo, permitiendo culminar positivamente con la implementación del módulo para la visualización de datos de producción científica de docentes investigadores en la Universidad Técnica de Cotopaxi a través de mapeo auto-organizado.

- **Estrategias comparativas:**

Es importante tener una estrategia para dar a conocer el producto que se realiza. En muchos casos el producto ofrecido es mejor que productos existentes y consolidados, es así que se debe explotar las funcionalidades del software. En el caso del Sistema Web ECUCIENCIA el módulo de visualización de datos de producción científica de docentes investigadores en la Universidad Técnica de Cotopaxi, es muy beneficioso ya que permite tener una ventaja considerable con sistemas que ofrecen servicios similares, además de las funcionalidades interactivas para el usuario orientado a utilizar el sistema.

- **Estrategias de adopción tecnológica:**

En los últimos años tanto como empresas como softwares se encuentran estables y dominantes en el mercado, es por esto importante las estrategias de adopción tecnológica para pensar en competir con sistemas similares, por lo cual es importante definir un ares, y a partir de esta identificar las acciones estratégicas en los diferentes contextos tecnológicos implicados. Para el sistema web ECUCIENCIA en general el desarrollo de este módulo para la visualización de información mediante mapeo auto-organizado es una de sus estrategias con la cual pretende causar impacto hacia los usuarios a los cuales está destinado este sistema.

En relación al módulo que se plantea, como principal estrategia se propone además de la visualización de los datos de producción científica de los docentes investigadores en la Universidad Técnica de Cotopaxi mediante mapeo auto-

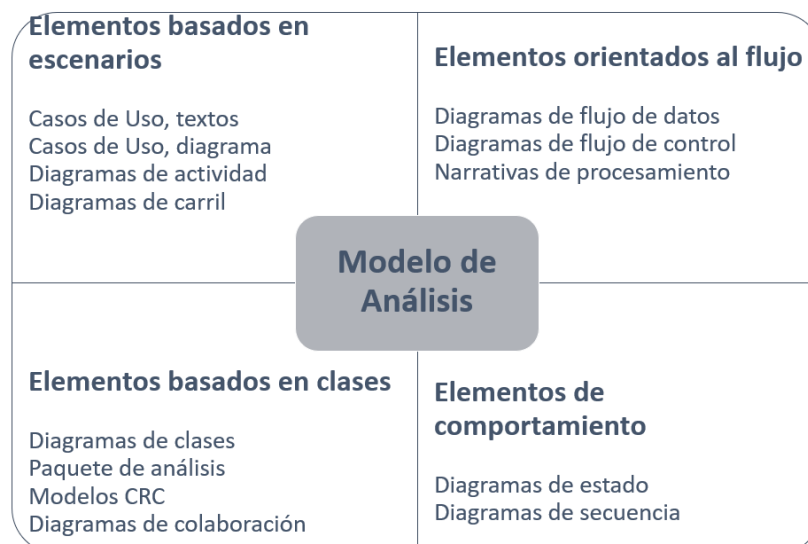
organizado, la capacidad de autoaprendizaje y toma de decisiones dándole un impulso tecnológico relevante, esperando causar un gran impacto y sobre todo brindar a los usuarios finales más que un sistema una herramienta tecnológica beneficiosa en cuanto a la actualizada que vivimos y la actualidad tecnológica que estamos ingresando, siendo en esta el principal punto de estudio la inteligencia artificial y toma de decisiones.

2.2.4. MODELOS DE ANÁLISIS, DISEÑO, IMPLEMENTACIÓN Y PRUEBA DE SISTEMAS DE INFORMACIÓN.

- **Modelo de análisis de sistemas de información:**

El modelo de análisis en los sistemas de información es la primera etapa durante el desarrollo de un sistema. La cual está estructurada por una serie de documentos y diagramas que representan los requisitos para el software y así definir las funcionalidades y comportamiento del mismo, facilitando el proceso de comprensión y aportando criterios desde distintos puntos de vista, con lo que disminuye la posibilidad de errores y descuidos en las siguientes etapas. Con el fin de cumplir con lo mencionado en el análisis se toma en cuenta los elementos mostrados en la Figura 2 a continuación.

Figura 18. Elementos del Modelo de Análisis



Fuente: [52]

De acuerdo a la Figura 20 propuesta por [52] los elementos del modelo de análisis son los enlistados a continuación y cada elemento está compuesto por la creación de escenarios o casos:

Elementos Basados en Escenarios: Para la mejor interacción con el sistema y el usuario:

- ✓ Casos de Uso.
- ✓ Diagramas de actividad.
- ✓ Diagrama de carril.

Elementos Orientados al Flujo: Tiene una visión del sistema del tipo entrada-proceso-salida:

- ✓ Diagrama de flujo de datos.
- ✓ Diagramas de flujo de control.
- ✓ Narrativas del procesamiento.

Elementos Basados en Clases: Una clase orientada a objetos encapsula atributos de los datos e incorpora las operaciones implicados en dichos atributos:

- ✓ Modelo CRC (clase-responsabilidad-colaborador).
- ✓ Diagrama de clases.
- ✓ Paquetes de análisis.
- ✓ Diagramas de colaboración.

Elementos de Comportamiento: Indica la forma en que el software responderá a los eventos o estímulos externos:

- ✓ Diagrama de estado.
- ✓ Diagrama de secuencia.

Tomando en cuenta los distintos modelos de análisis y la utilidad que estos proponen para el desarrollo de un software, durante el desarrollo del módulo de visualización de datos de producción científica de docentes investigadores en la Universidad Técnica de Cotopaxi para el sistema web ECUCIENCIA, de acuerdo a los escenarios se construirán los siguientes diagramas en el software Visual Paradigm, los mismos que impulsarán y aportarán positivamente durante el desarrollo:

- ✓ Diagrama de casos de uso.
- ✓ Diagrama de actividad.
- ✓ Diagrama de clases.
- ✓ Diagrama de estado.
- ✓ Diagrama de secuencia.

- **Modelo de diseño de sistemas de información:**

La siguiente etapa en el desarrollo del sistema es el diseño, la cual está considerada para las opciones arquitectónicas, es decir en esta etapa nos encargamos de la elaboración del boceto de las interfaces que serán a continuación codificadas. Estos bocetos deben ser elaborados a partir del modelo de análisis de sistemas de información anteriormente mencionados y analizados, de los mismos deben tomarse en cuenta las diferentes etapas, tomando como principales los casos de uso, diagramas de flujo, de actividad y los que crea pertinentes, en las cuales se detalla los requerimientos en si del software a desarrollarse.

Para el diseño del módulo de visualización de datos de producción científica de docentes investigadores en la Universidad Técnica de Cotopaxi mediante mapeo auto-organizado para el sistema web ECUCIENCIA se utilizará como herramienta de diseño Balsamiq Mockups, el mismo que permitirá dar el boceto adecuado para posteriormente cuando lo definamos sea puesto en consideración durante el proceso de implementación y codificación del mismo. Además, el software mencionado ofrece muchas herramientas de fácil uso para realizar dichos bocetos fácil y rápidamente, siendo de gran ayuda para el correcto desarrollo y la pronta elaboración del módulo.

- **Modelo de implementación de sistemas de información:**

El modelo de implementación de sistemas de información es principalmente la codificación o programación del sistema a desarrollarse, el cual consiste en llevar a código fuente de acuerdo a un lenguaje de programación elegido por el programador, todo lo diseñado en la fase anterior. El programador debe seguir por completo lo sugerido en la fase de diseño y la fase de análisis. Se puede considerar que en esta etapa se consume la mayor cantidad de tiempo, pero es esencialmente

indispensable el correcto análisis y diseño para que esta etapa sea realizada de forma concisa, rápida y esencialmente efectiva.

Además, se considera el tiempo tomado en la depuración del código por errores suscitados durante la fase de desarrollo, aunque la siguiente etapa se encarga de encontrar posibles bugs para poder liberar la solución presentada por el programador, es importante desde este punto entregar el software con la menor cantidad de bugs posibles. Durante esta fase, el código puede adoptar varios estados dependiendo de la forma de trabajo y el lenguaje de programación utilizado, son así los casos directamente de codificación, casos referentes al compilador y casos concernientes al ejecutable.

Para la implementación (codificación) del módulo de visualización de datos de producción científica de docentes investigadores de la Universidad Técnica de Cotopaxi mediante mapeo auto-organizado para el sistema web ECUCIENCIA se utilizarán las siguientes herramientas:

- ✓ Lenguaje de programación Python.
- ✓ Framework Django.
- ✓ Librería SOMPY.
- ✓ Base de Datos PostgreSQL.
- **Modelo de prueba de sistemas de información:**

En este modelo de sistemas de información se efectuarán el testeo del sistema en desarrollo, principalmente las pruebas unitarias y de integración. Estas pruebas se los realiza con un conjunto seleccionado de datos típicos a los que puede verse sometido el sistema, además las conocidas pruebas de estrés en las cuales se lleva al límite la respuesta del software con el fin de probar su tolerancia y robustez. También se debe tomar en cuenta pruebas referentes a situaciones en que normalmente el sistema no va estar sometido o no debería estarlo, pero suele suceder por error o casualidad.

Generalmente suele existir pruebas finales del software conocidas como Beta Test, en el cual el sistema es instalado en condiciones parecidas al que se va encontrar a partir de su liberación, esto con el fin de encontrar posibles errores, inestabilidades, entre otros, los cuales pudieron escapar de las pruebas anteriores. Esta prueba

generalmente es realizada por personas conocedoras del funcionamiento final del sistema, normalmente estos son los usuarios finales los cuales pactan un periodo de testeo con el desarrollador. En caso de encontrarse con errores o fallos son transmitidos a los programadores para su depuración.

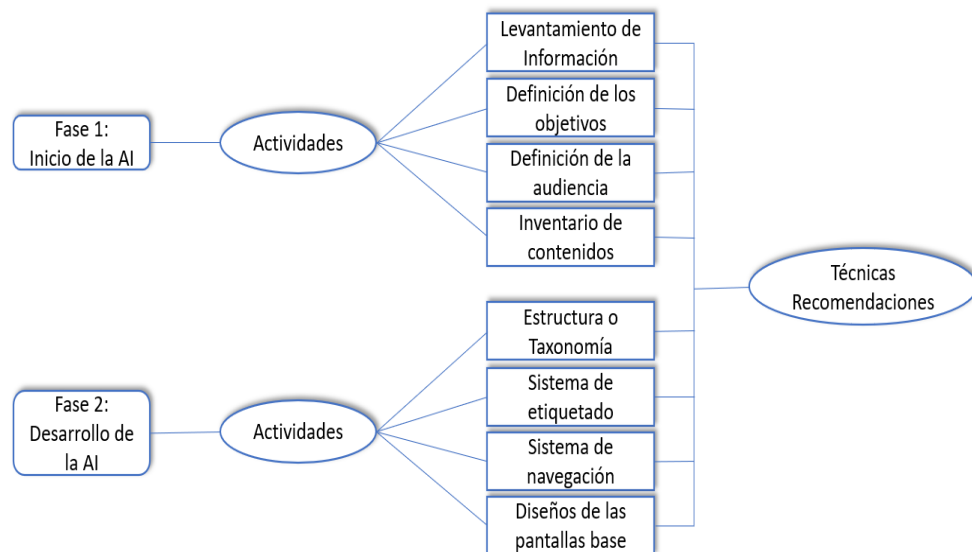
En el caso del módulo propuesto de visualización de datos de producción científica de docentes investigadores de la Universidad Técnica de Cotopaxi mediante mapeo auto-organizado para el sistema web ECUCIENCIA, se realizará las pruebas sugeridas para no tener problemas o fallos a partir de su liberación. Estas pruebas serán realizadas y seguidas rigurosamente por parte del desarrollador y así utilizar el cien por ciento de sus funcionamientos dando el empujón necesario al sistema web mencionado, con lo cual se pretende sobresalga en comparación de sistemas que ofrecen servicios similares.

2.2.5. METODOLOGÍA DE ARQUITECTURA DE INFORMACIÓN Y TOMA DE DECISIONES:

- **Metodología de arquitectura de información:**

La arquitectura de información pretende organizar grandes cantidades de contenidos con la finalidad de que el usuario pueda utilizarla de forma efectiva y sencilla. [53] en el año 2012 proponen una guía práctica, la cual supone será una herramienta de ayuda y soporte en el proceso de elaboración de arquitecturas de información, en las que se define las actividades, técnicas y recomendaciones aplicables en cada etapa del proceso. Se estructura en dos fases principales que se pueden observar en la Figura 21 a continuación presentada.

Figura 19. Flujo de actividades de las dos fases de guía práctica.



Fuente: [53]

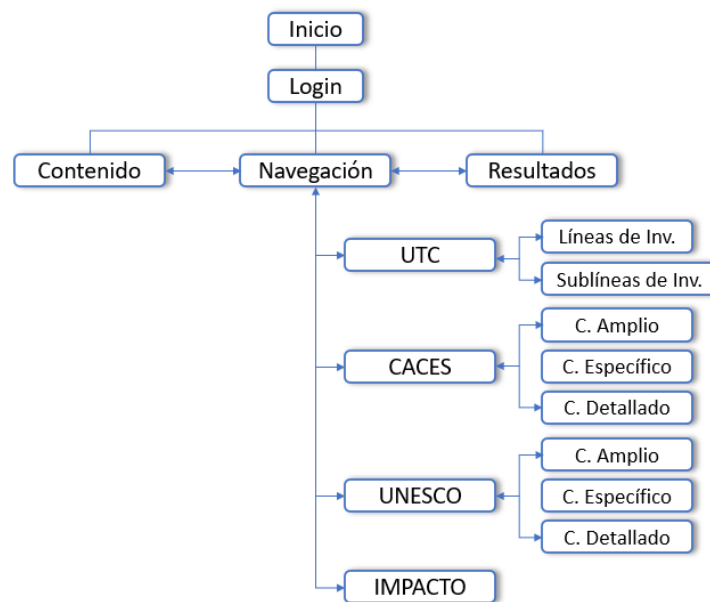
Fase 1: Inicio de la Arquitectura de información.

- **Recopilación de información:** Las técnicas utilizadas son las entrevistas y encuestas para conocer las necesidades del usuario. Se recomienda involucrarse el proyecto desde el inicio, tener en cuenta las metodologías de desarrollo a utilizar, tomar en cuenta la participación directa del usuario, estudio de la audiencia y un análisis competitivo de las tecnologías y herramientas de productos similares al pretendido.
- **Definición de los objetivos:** Las preguntas deben ser enfocadas al propósito de la aplicación. Se recomienda identificar requisitos funcionales para conocer a fondo lo que se pretende realizar.
- **Definición de la audiencia:** Se debe estudiar a los usuarios, para lo cual se puede utilizar encuestas, entrevistas y técnicas de card sorting. Se recomienda utilizar metodologías existentes para estos casos, conocer el ámbito en el que se pretende implementar la solución y entrevistarse con el o los usuarios involucrados con las necesidades.
- **Inventario de contenidos:** se debe utilizar las entrevistas. Se recomienda ser exhaustivo durante las entrevistas para obtener contenidos coherentes para el desarrollo.

Fase 2: Desarrollo de la Arquitectura de Información.

- **Estructura o taxonomía:** Se recomienda tener los requisitos funcionales del producto, tomar en cuenta el orden lógico para la presentación de los contenidos y se recomienda utilizar un modelo jerárquico simple.
- **Sistema de etiquetado:** Se debe definir las etiquetas que tendrá la aplicación. Se recomienda definir el sistema de etiquetado basándose en los objetivos del usuario, concentrarse en el público objetivo (docentes investigadores), colocar en orden las etiquetas y con nomenclatura coherente.
- **Sistema de navegación:** Se recomienda continuar con este punto solo y únicamente si ya tiene los puntos anteriores realizados, permitir navegar con facilidad y agilidad dentro del contenido y crear el mapa de navegación para definir como quedara el producto final como se muestra en la Figura 22.

Figura 20. Mapa de navegación del módulo con el método SOM



Elaborado por: Rivera Chasiquiza Alex Geovanny

- **Diseño de las pantallas base:** La técnica a utilizar será la entrevista para definir las características visuales del producto. Se recomienda reunirse con el usuario para definir las características visuales, diseñar las pantallas base para lo cual se recomienda utilizar el diagrama de presentación, diseñar tantos diagramas como sean necesarios, reunirse con el usuario para presentarle las interfaces diseñadas obteniendo sugerencias para mejoras.
- **Toma de decisiones:**

Los sistemas que incluyen como apoyo la toma de decisiones ayudan a priorizar objetivos, evaluar alternativas y simular resultados, los mismos que serán interpretados y presentados para el usuario. Estos resultados serán de gran ayuda, por ejemplo, en un hospital especializado en tratar el cáncer, al obtener los resultados desde un sistema que procesa los datos, evalúa alternativas y simula resultados ayudara al médico encargado en el tratamiento incrementando las posibilidades de una posible cura. Para que un sistema sea capaz de interpretar los datos y tomar decisiones a partir de estos es necesario incluir inteligencia artificial en el mismo, es así que se pretende con el módulo propuesto sea esto posible en el sistema web ECUCIENCIA.

2.2.6. ESTÁNDARES, AMBIENTES Y HERRAMIENTAS DE SOFTWARE POR UTILIZAR:

En la propuesta de investigación se plantea desarrollar un módulo de visualización de datos de producción científica de docentes investigadores en la Universidad Técnica de Cotopaxi mediante mapeo auto-organizado, para lo cual es sumamente importante la correcta elaboración principalmente de la metodología a utilizarse, en el caso de esta investigación será utilizada la metodología iterativa incremental debido a que permite organizar el desarrollo por etapas funcionales. Con respecto al análisis de requerimientos, se puede evidenciar es la etapa de partida y por tanto es indispensable sea realizada correctamente ya que de esta depende el resto de etapas, es por eso importante el desarrollo de los diferentes diagramas recomendados en esta etapa.

Se recomienda seguir al pie la metodología seleccionada, como se mencionó en los modelos para el desarrollo del sistema web los diagramas a utilizarse son: los diagramas de casos de uso, diagramas de actividad, diagramas de flujo de datos, diagramas de clases, diagramas de estado y diagramas de secuencia. Estos fueron seleccionados por el aporte que proporcionan a partir de la concepción de los requisitos hasta la liberación de la solución. Además, se el sistema será desarrollado en el lenguaje de programación Python, framework Django y base de datos PostgreSQL utilizando la librería SOMPY la cual permite procesar los datos mediante mapeo auto-organizado.

Una vez definidos los procesos, procedimientos y herramientas de software a utilizar se estima el desarrollo de la propuesta de investigación es factible, ya que aportará al sistema web ECUCIENCIA grandes beneficios en cuanto a funcionalidades con respecto a sistemas con procesos similares y a la Universidad Técnica en general ya que incursionara en un tema muy importante que en la actualidad está tomando gran importancia como es la Inteligencia Artificial (IA) y toma de decisiones. Además, que esta investigación será el punto de partida para que más investigadores se integren y mejoren el tema en cuestión para el beneficio DE LA UNIVERSIDAD TÉCNICA DE COTOPAXI.

2.3. DISEÑO EXPERIMENTAL Y MÉTODO DE CRITERIO DE EXPERTOS

2.3.1. DISEÑO EXPERIMENTAL

En la Universidad Técnica de Cotopaxi se pretende realizar un módulo encargado de la visualización de información de datos de producción científica producida por docentes investigadores, esta visualización se pretende sea expresada al usuario mediante mapas de calor al cumplir una serie de etapas desde la primera clasificación, interactuando directamente con la data existente en el sistema ECUCIENCIA. El proceso de automatización de lo planteado será realizado en coordinación con el usuario responsable de dicho sistema, permitiendo la resolución a la problemática en la visualización de grandes cantidades de datos y la toma de decisiones de acuerdo al nivel de impacto de artículos en revistas según su prestigio a nivel académico e internacional.

Habiendo definido el proceso a realizarse y experimentar correctamente, se muestra a continuación el diseño de las fases a tomarse en el desarrollo del módulo planteado:

- **Comprensión y planeamiento del problema:**

Resulta complicado realizar este proceso ya que aparenta ser sencillo el plantear un problema es completamente lo opuesto. En este proceso es importante interactuar con todas las partes involucradas. Para estar claros en esta etapa se debe tener el mayor conocimiento posible referente a lo que se busca expresar y solucionar. En

cuando al sistema ECUCIENCIA se tomará en cuenta los requerimientos obtenidos mediante la observación y entrevista realizada al responsable directo.

- **Elección de factores y niveles:**

En esta parte de la experimentación es necesario el experimentador tenga la experiencia y comprensión teórica necesaria para no trabarse y poder dar una solución real y pronta. En la Universidad Técnica de Cotopaxi el problema expresado es claro, lo que se debe tomar en cuenta es el manejo de datos que están realizando para la preclasificación de los mismos.

- **Selección de la variable de respuesta:**

La variable de respuesta debe representar la solución concreta de la problemática investigada. Para el sistema ECUCIENCIA la respuesta será la visualización mediante mapeo auto-organizado de los datos de producción científica de artículos, libros, ponencias y revistas según las líneas y sublíneas de investigación en la Universidad Técnica de Cotopaxi y el campo amplio, específico y detallado de acuerdo a la UNESCO y CACES.

- **Elección del diseño experimental:**

Una vez definidas correctamente las fases expuestas anteriormente, es importante la correcta elección del diseño experimental, para lo cual se toma en cuenta el tamaño de la muestra y enfocarse en los resultados esperados. En el caso de la Universidad Técnica de Cotopaxi se cuenta con 470 docentes investigadores, los mismos que serán tomados en cuenta.

- **Realización del experimento:**

Durante la puesta en marcha del experimento es necesario monitorear los procesos y resultados que se vayan obteniendo. En el caso del módulo a realizarse se seguirá como lineamiento el modelo iterativo incremental para mejores resultados.

- **Análisis de resultados:**

Para el análisis de resultados es prudente enfocarse en lo planteado desde la recolección de información y sus procesos, de esta manera se podrá medir el nivel

de satisfacción en la visualización mediante mapeo auto-organizado de los datos de producción científica en la Universidad Técnica de Cotopaxi. Algo adicional que se debe tomar en cuenta es el cronograma planteado para la experimentación, la cual debe seguir lo más estrictamente posible.

2.3.2. MÉTODO DE CRITERIO DE EXPERTOS

- **Método Delphi**

Este método es ampliamente utilizado para la correcta validación de propuestas científicas planteadas por investigadores, la cual consiste en la aplicación sucesiva de cuestionarios a un grupo de expertos, los cuales son seleccionados adecuadamente por sus conocimientos y habilidades referentes a la investigación que se encuentra en desarrollo. Es necesario entre cada ronda analizar estadísticamente los resultados obtenidos a partir de los cuales se genera un nuevo cuestionario que deberá ser redistribuidos a los expertos, esto se lo realiza hasta lograr obtener una estabilidad en los resultados estadísticos. Para mejores resultados estudiosos recomiendan calcular el estadístico de fiabilidad, Alfa de Cronbach o índice de consistencia interna, con la cual se estima la precisión entre las preguntas del cuestionario planteado. EL Método Delphi está compuesto por las siguientes Fases:

Fase preliminar

Consiste en formar grupos que permitan implementar adecuadamente el método en cuestión, por lo que se conformó dos grupos. El grupo coordinador está encargado de desarrollar la investigación, selección del grupo de expertos, validación e implementación de los cuestionarios en las diferentes etapas e interpretar los resultados estadísticos parciales y finales. El segundo grupo es el de expertos, en el cual sus integrantes se encargan de llenar las encuestas en cada etapa necesaria de manera anónima y concisa acorde a las preguntas planteadas.

Selección de expertos

Los integrantes del grupo de expertos son sumamente importantes para la obtención de resultados fiables, es por esto necesario que dichos expertos sean seleccionados acordes a su profesión, cargo, años de experiencia, categoría docente, grado

científico, conocimiento de la investigación, entre otros. Ciertos estudiosos en el tema hacen mención que el mínimo de expertos debe ser siete y el máximo no superior a treinta.

Fase exploratoria

En esta fase mediante la extracción de criterios y atributos a evaluar se elabora y ejecuta las diferentes etapas de los cuestionarios. Es importante recalcar que las preguntas deben ser concisas y acorde a la investigación planteada, adicional, para facilitar el trabajo de los expertos se debe acompañar cada criterio del análisis estadístico realizado en la siguiente versión del cuestionario y es recomendable incluir una pregunta abierta para recolectar valoraciones cualitativas referentes a los criterios planteados. Además, las etapas deberán ser repetidas las veces necesarias hasta perfeccionar el cuestionario final, en el cual se obtendrán los resultados que ayudarán a dar solución a la problemática.

Fase final

La fase final se encarga de sintetizar los resultados obtenidos durante el proceso de validación con el grupo de expertos, esto de acuerdo a la versión final del cuestionario. Los mismos que permitirán validar la investigación referente a la clasificación y visualización de información cuantitativa de docentes investigadores en la Universidad Técnica de Cotopaxi mediante SOM de sus líneas y sublíneas de investigación, y de los campos amplios, específicos y detallados de acuerdo a la CACES y UNESCO.

Resultados

Al culminar con el análisis estadístico de la versión final del cuestionario se obtienen los resultados con su valoración, con la cual se valida la investigación. Para que estos resultados sean confiables es importante la selección de los expertos como se mencionó en las fases anteriores, por lo que es necesario obtener el coeficiente de competencia de los expertos.

2.4. DESCRIPCIÓN METODOLÓGICA DE LA VALORIZACIÓN ECONÓMICA, TECNOLÓGICA, OPERACIONAL Y MEDIO AMBIENTAL DE LA PROPUESTA.

2.4.1. VALORACIÓN ECONÓMICA:

En la valoración económica es importante mencionar los costos directos e indirectos con los cuales se obtendrá los costos reales de la propuesta de investigación. En cuanto al desarrollo de software se prefirió utilizar herramientas open source, de esta manera disminuir considerablemente costos al culminar el proyecto.

2.4.2. VALORACIÓN TECNOLÓGICA:

En esta valoración se destaca la creación de un módulo, el cual será el encargado de clasificar, aprender y visualizar la información existente en la base de datos del sistema ECUCIENCIA. En este contexto fue necesario la creación de una metodología acorde a las necesidades presentadas en el transcurso de la investigación. La metodología mencionada es la de mapeo auto-organizado (SOM), la cual tiene por objetivo la correcta interacción y clasificación de la data para su procesamiento y obtención de resultados fiables.

2.4.3. VALORACIÓN AMBIENTAL:

Al ser el ECUCIENCIA un sistema web, los docentes investigadores en calidad de usuarios del mismo, se los incentiva a digitalizar sus artículos, libros y ponencias teniendo la posibilidad de acceder desde cualquier lugar, con lo que disminuye considerablemente el uso de archivos físicos para el traslado de la información.

2.5. CONCLUSIONES CAPITULO II

- Las metodologías planteadas para el proceso de investigación son importantes, ya que facilitan la interacción desde la recolección de información hasta la obtención de resultados reales de la problemática a resolver.
- El uso de la metodología de mapeo auto-organizado (SOM) será beneficiosa en la resolución del problema planteado durante la investigación, debido a que permitirá analizar la data del sistema ECUCIENCIA para su correcto aprendizaje y visualización de resultados para el usuario.
- La implementación tanto de la metodología como del módulo SOM en el sistema ECUCIENCIA presentara beneficios directa e indirectamente al manejo de datos y su procesamiento, lo que facilitara la creación e implementación de las redes neuronales artificiales y la visualización de acuerdo a los parámetros establecidos por la UNESCO, CACES y Universidad Técnica de Cotopaxi.

CAPÍTULO III

3. APLICACIÓN Y/O VALIDACION DE LA PROPUESTA

3.1. RESULTADOS DE DIAGNÓSTICO DEL PROBLEMA

La investigación está enfocada a la visualización de información de datos de producción científica en la Universidad Técnica de Cotopaxi mediante mapeo auto-organizado, la cual será implementada en un nuevo módulo que se encargará de clasificar los procesos conforme a las áreas del conocimiento, colaboración de investigadores como artículos, libros y ponencias; además de acuerdo a las líneas y sublíneas de investigación en la Universidad Técnica de Cotopaxi; y los campos amplios, específicos y detallados acorde a la UNESCO y CACES. Con las necesidades planteadas adecuadamente para el mejoramiento del sistema ECUCIENCIA se determinó factible el desarrollo e implementación del método de trabajo y módulo ya que brindara beneficios precisos conforme la problemática.

Como se menciona durante el proceso de investigación para dar solución a la problemática hallado en el proceso de producción científica en la Universidad Técnica de Cotopaxi, fue necesario la creación de un método de trabajo alineado a SOM que permita solucionar de manera precisa el manejo de la data su aprendizaje y visualización. EL método de trabajo en cuestión es el de mapeo auto-organizado (SOM) que tienen la capacidad de procesar los datos del sistema ECUCIENCIA para mediante el uso de redes neuronales artificiales se proceda con el entrenamiento, agrupamiento, validación y visualización de resultados acordes a la data de carga o data inicial. Este resultado obtenido debe ser mostrado al usuario en un mapa de calor debido a su contexto y para la mejor interpretación de la misma. Cabe recalcar que el SOM es una técnica de IA enfocada en la clasificación no supervisada de información con lo cual el sistema interactuara directamente con la

base de datos y el usuario sin la necesidad de terceros que se encarguen del proceso. Esto se realiza con el objetivo de dar mayor solvencia y mejorar la automatización de los procesos de producción científica mencionados en la Universidad Técnica de Cotopaxi y ayudar al sistema ECUCIENCIA en sus resultados esperados.

3.1.1. TÉCNICAS DE INVESTIGACIÓN

- **Observación**

Al interactuar con el sistema ECUCIENCIA se pudo comprobar su funcionalidad, en el cual se dio a notar la particularidad que existe una falencia en la clasificación y visualización de datos de producción científica referente a docentes investigadores en la Universidad Técnica de Cotopaxi, además, referido a la clasificación de la información a ser mostrada tiene que ser procesada montando servicios en Windows que se ejecutan cada cierto periodo de tiempo, lo que no garantiza una información inmediata al momento de visualizar los resultados.

- **Entrevista:**

Las entrevistas fueron realizadas al PhD. Gustavo Rodríguez, que para ese entonces brindaba el servicio de coordinador del proyecto relacionado a la Red de Estudios Cienciométricos (REDEC), el cual puso en marcha el desarrollo del sistema ECUCIENCIA. Con las respuestas dadas en dicha entrevista se procede al análisis de requerimientos para el módulo de visualización de información cienciométrica mediante mapeo auto-organizado. La información a procesar deberá ser sometida a técnicas metodológicas para sus resultados sean óptimos. La visualización de datos de producción científica de los investigadores de la Universidad Técnica de Cotopaxi mediante mapeo auto-organizado, será clasificada mediante las líneas de investigación de acuerdo a la Universidad Técnica de Cotopaxi; y de acuerdo a los campos amplios, específicos y detallados según criterios del CACES y UNESCO.

Por lo expresado, durante la entrevista se vio prudente realizar el siguiente cuestionamiento a los encargados:

Tabla 3. Análisis de la Entrevista

N°	PREGUNTA	RESPUESTA	CONCLUSIÓN
1	¿De qué trata el sistema ECUCIENCIA?	Actualmente se ha venido desarrollando el sistema como proyectos de pregrado con alumnos de la carrera de Ingeniería en Informática, el sistema ECUCIENCIA trata sobre el manejo de datos cuantitativos de los docentes investigadores de la Universidad Técnica de Cotopaxi, es decir de sus artículos, libros o ponencias que van desarrollando y agregando a revistas científicas.	El sistema ECUCIENCIA es un proyecto ambicioso que intenta manejar grandes cantidades de datos, los cuales son ingresados por los docentes investigadores de la Universidad Técnica de Cotopaxi en primera instancia, pero teniendo la intención de expandirse a las demás Universidades.
2	¿Qué información gestiona el sistema ECUCIENCIA?	La información que gestiona es como ya se mencionó los artículos, libros y ponencias que son vinculadas a revistas científicas, lo que ayuda a los docentes de la Universidad Técnica de Cotopaxi en la toma de decisiones, la cual es mostrada en el sistema para su análisis.	La información que gestiona el sistema ECUCIENCIA son artículos, libros y ponencias que son vinculadas en revistas de diferentes categorías y de esta manera tener un control de esto.
3	¿Quiénes son los usuarios que interactúan directa o indirectamente con el sistema ECUCIENCIA?	Los usuarios que interactúan directamente con el sistema son los docentes investigadores que registran sus artículos, libros y ponencias en el sistema. Además, al ser un sistema web existen usuarios externos que tendrán acceso a la visualización de los resultados.	En primera instancia los usuarios directos son los docentes investigadores, adicional están los administradores y usuarios en general que podrán apreciar lo que el Sistema ECUCIENCIA propone.
4	¿Cuál es la información necesaria para el mapeo auto-organizado?	Para el mapeo auto-organizado es necesario analizar los datos de artículos, libros y ponencias que se han registrado en el sistema, aquí se debe tomar en cuenta las líneas y sublíneas de investigación que rigen en la Universidad Técnica de Cotopaxi; y los campos	Ya que el proyecto en general trata sobre el estudio y manejo proyectos de docentes investigadores anexados a revistas de diferentes categorías, la data a ser tratada será acorde a los artículos, libros y ponencias que los

		amplios, específicos y detallados de acuerdo a la UNESCO y CACES.	docentes investigadores registran en el sistema ECUCIENCIA. Dicha información está estructurada de acuerdo a las líneas y sublíneas de Investigación en la Universidad Técnica de Cotopaxi; además de los campos amplios, específicos y detallados de acuerdo al CACES y UNESCO.
5	¿La información se encuentra clasificada correctamente para su manejo?	La información es registrada por los docentes investigadores directamente en el sistema por lo que estaría clasificada de acuerdo a sus criterios. En este sentido está clasificada por la Facultad y Carrera a la que el docente pertenece.	La data se encuentra estructurada acorde a la información del docente investigador, es decir a la facultad y carrera que dicho docente pertenece.
6	¿En qué lenguaje de programación fue desarrollado el sistema ECUCIENCIA y utiliza algún framework?	El lenguaje de programación con el que se ha venido desarrollando el sistema es Python, este se lo ha utilizado por su manejo de datos y librerías existentes para el desarrollo de proyectos de este tipo. Al ser un sistema web se optó por la utilización del framework Django.	Actualmente python es un lenguaje muy robusto para el manejo de grandes cantidades de datos. Por tal motivo es prudente su utilización para el desarrollo de esta propuesta planteada.
7	¿Qué gestor de base de datos es con el que se ha venido trabajando?	El sistema se ha desarrollado con el gestor de base de datos PostgreSQL, debido a su robustez en el manejo de datos y excelente acoplamiento con el framework y lenguaje de programación utilizado.	Para el manejo y almacenamiento de información existen varios gestores, PostgreSQL es una buena opción para el sistema ECUCIENCIA ya que es muy robusta para el manejo de grandes cantidades de datos.

Elaborado por: Alex Geovanny Rivera Chasiqiza

3.2. 3.2. RESULTADOS DE LOS MÉTODOS ESPECÍFICOS

Los métodos específicos son empleados para la recolección de información permitiendo continuar con el desarrollo de la propuesta planteada de visualización de datos de producción científica relacionada a docentes investigadores en la Universidad Técnica de Cotopaxi mediante mapeo auto-organizados. El método de trabajo para la investigación planteada como ya se mencionó está alineada al mapeo auto-organizado (SOM), el cual fue diseñado acorde a las necesidades recolectadas para dar una solución fiable a la problemática existente. En cuanto a la metodología de desarrollo se decidió por la iterativa incremental, ya que este permite obtener prototipos funcionales en un tiempo determinado de acuerdo al cronograma partiendo desde el análisis de requerimientos, diseño, implementación, pruebas y la puesta en funcionamiento. A continuación, se expresan los resultados obtenidos en cuanto a las técnicas de investigación y metodologías utilizadas en el proceso de dar solución a lo propuesto.

Como resultado de la entrevista realizada a los coordinadores de REDEC se obtuvo los requerimientos necesarios para la puesta en marcha de la propuesta de investigación. En la entrevista se obtuvo las problemáticas que tiene actualmente el sistema ECUCIENCIA, para lo cual es pertinente el uso de la metodología planteada para la correcta clasificación de los datos de producción científica en la Universidad Técnica de Cotopaxi de acuerdo a sus líneas y sublíneas de investigación; y conforme los campos amplios, específicos y detallados que la UNESCO y CACES proponen respectivamente. Los requisitos encontrados con la utilización de las técnicas planteadas serán el punto de partida para la correcta funcionalidad e interacción entre el usuario y el sistema.

Los requerimientos solicitados avalarán las funcionalidades del módulo y del sistema ECUCIENCIA en general, evitando interferir en el proceso de subida de información de los docentes investigadores. Además, es esencial de la propuesta que usuarios externos puedan interactuar fluidamente con la información mediante el mapeo auto-organizado garantizando la información sea verídica, también de la fácil interacción y funcionalidad de la misma.

3.2.1. MÉTODO DE TRABAJO ALINEADO A MAPEO AUTO-ORGANIZADO (SOM)

El mapeo auto-organizado (SOM) corresponde a la clasificación y visualización de información utilizando redes neuronales artificiales para el aprendizaje no supervisado y análisis de los datos. Tomando como referencia investigaciones realizadas respecto a este tema en distintas partes del planeta, a pesar de ofrecer muchos beneficios para los investigadores, no se le ha puesto el suficiente énfasis. Esto empezó a cambiar durante los últimos años donde se ha evidenciado el manejo de grandes cantidades de datos y con ello la búsqueda de una metodología estable que permita analizarlos e interpretarlos.

Bajo esta compleja necesidad de encontrar una metodología que enmarque la propuesta de investigación planteada se vio prudente el diseño e implementación del mapeo auto-organizado (SOM) como método de trabajo, con la cual se pretende analizar de manera clara y efectiva las etapas que se deben tener en cuenta para el correcto análisis e interpretación de resultados desde la carga de datos a las redes neuronales artificiales, hasta la visualización mediante mapas de calor. En la Figura 13 se describe el método de trabajo alineado a SOM interpretada gráficamente, en la cual se puede comprobar los procesos de carga de datos, entrenamiento, agrupamiento, validación y visualización todo esto partiendo de una base de datos acoplada para la obtención de resultados por lo que antes de interactuar con el método de trabajo es necesario un análisis previo y clasificación de los datos a procesar.

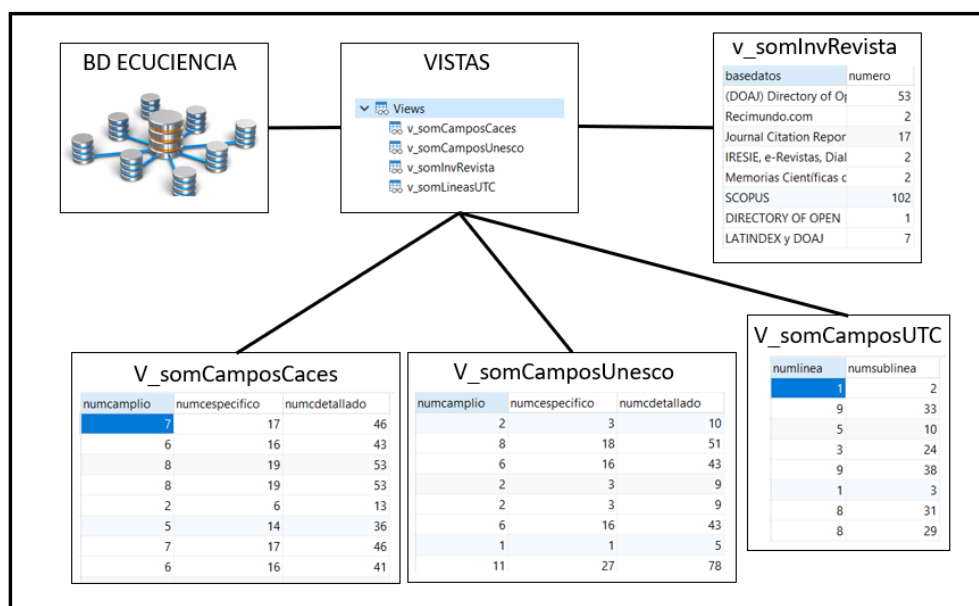
La estructura del método de trabajo alineado a SOM propuesta será descrita a continuación, tomando en cuenta la investigación realizada para dar solución a la problemática en procesar la producción científica de docentes investigadores en la Universidad Técnica de Cotopaxi de acuerdo a las líneas y sublíneas de investigación; y los campos amplios, específicos y detallados según la CACES y UNESCO:

ETAPA 1. Carga de datos:

Hace referencia a la data que será ingresada a la red neuronal artificial, esta data debe estar previamente analizada y clasificada acorde a las necesidades para el

adecuado entrenamiento. Con respecto a la investigación propuesta la data inicial se la coincidió con el análisis del funcionamiento y relaciones del sistema web ECUCIENCIA integrado a su base de datos, en la cual fue importante conocer las relaciones entre tablas y de esta manera crear vistas que ayudarán a la clasificación de la información, cuyos resultados serán ingresados a las redes neuronales artificiales para su entrenamiento. En la Figura 14 se muestra su interpretación y en la Figura 23 a continuación presentada se detalla el proceso de esta carga de datos en la implementación del método de trabajo alineado a SOM.

Figura 21. Carga de Datos



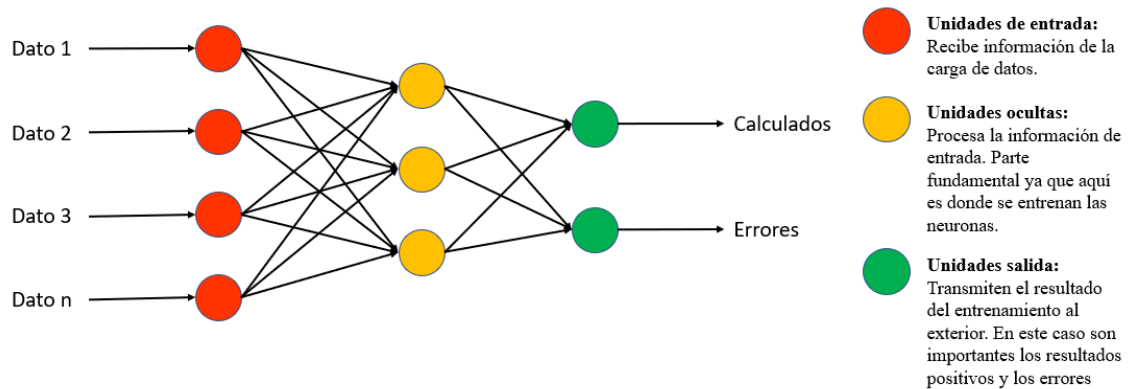
Elaborado por: Alex Geovanny Rivera Chasiquiza

ETAPA 2. Entrenamiento:

El entrenamiento es la parte fundamental de la método de trabajo alineado a SOM, debido a que este proceso es el encargado de recibir los datos en bruto y procesarlos para su aprendizaje, es decir mediante un algoritmo ingresa la data a la red neuronal y se encarga de leer y analizar cada dato para obtener resultados y predecir el comportamiento de los mismos, luego evalúa esos datos resultantes para su aprendizaje y por último se encarga de devolver la data entrenada para su siguiente paso dentro del método de trabajo. Para ser más específico en este proceso se muestra en la Figura 15 gráficamente como se realiza el proceso de entrenamiento y entrega de resultados. Además, en la Figura 24 a continuación presentada se

explica el proceso de entrenamiento realizado en el desarrollo del módulo para el sistema ECUCIENCIA.

Figura 22. Entrenamiento de las neuronas



Elaborado por: Alex Geovanny Rivera Chasiquiza

Los datos ingresados de las vistas mencionadas en la fase anterior corresponden a las líneas y sublíneas de investigación, y a los campos amplios, específicos y detallados con respecto al CACES y UNESCO. En las cuales se interactúa con las tablas responsables de almacenar sus datos. En el entrenamiento se pretende obtener resultados reales y confiables para poder presentarlos a la siguiente etapa encargada de la agrupación de los mismos. Para lo cual se utiliza la librería SOMPY que se encarga de entrenar la data inicial indicada en la figura 24 mediante RNA.

ETAPA 3. Agrupamiento:

En esta fase se reciben los datos resultantes del entrenamiento a las vistas que contenían la información de las tablas requeridas para su procesamiento. El agrupamiento consiste en ordenar de acuerdo al rango de resultados y al campo o línea que corresponde para una visualización más funcional y amigable con los usuarios. En la Figura 16 mostrada se puede apreciar de mejor manera en lo que consiste el agrupamiento y en la Figura 25 a continuación presentada se identifican la matriz resultante de los datos entrenados y agrupados para su posterior validación y visualización.

Figura 23. Agrupamiento de los datos resultantes



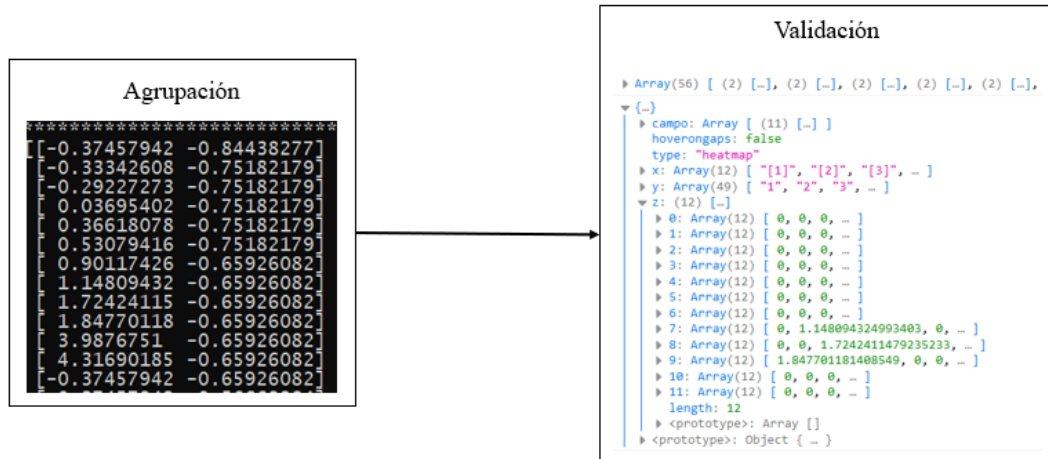
Elaborado por: Alex Geovanny Rivera Chasiquiza

Los datos resultantes del entrenamiento de acuerdo a las líneas y sublíneas de investigación de la Universidad Técnica de Cotopaxi y los campos amplios, específicos y detallados según el CACES y UNESCO serán agrupados por rangos para poder visualizarlos conforme lo explicado, una mejor interacción entre el sistema ECUCIENCIA y los usuarios.

ETAPA 4. Validación:

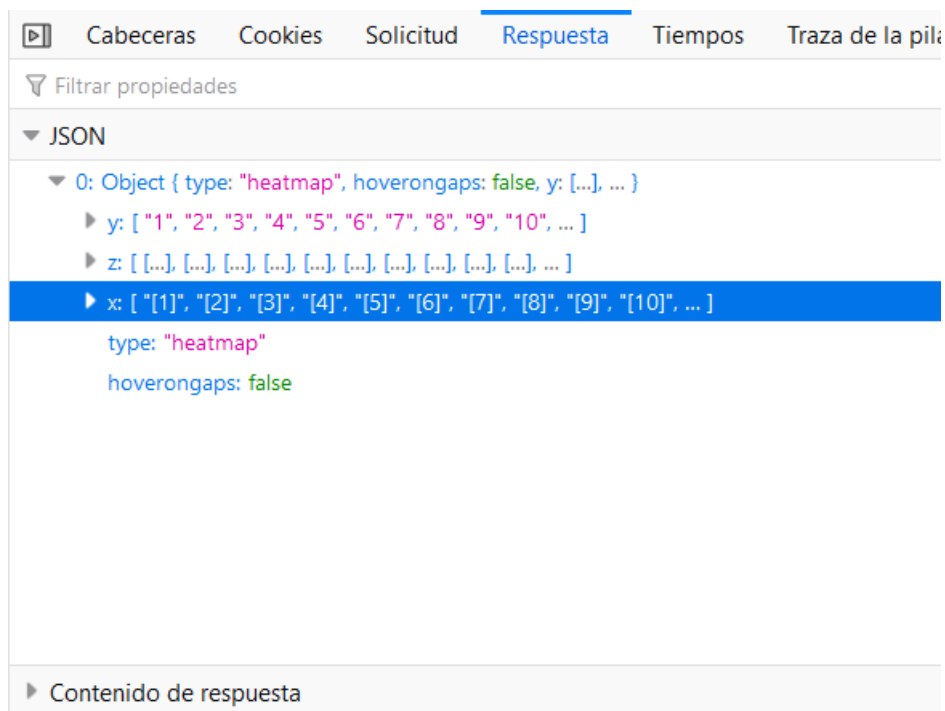
Con los datos ya agrupados es prudente validarlos para verificar no existan falencias en los mismos, con este paso nos aseguramos de la correcta visualización e interacción entre sistema y usuario. En la Figura 17 se explica gráficamente el significado de validación y en las Figuras 26 y 27 se muestra el resultado de la validación, que para la interpretación entre el Template y View del Framework Django se utilizó codificaciones en JavaScript, JQuery y Ajax respectivamente de manera asíncrona para evitar la pérdida de información al presentar los resultados gráficos.

Figura 24. Validación de la información



Elaborado por: Alex Geovanny Rivera Chasiquiza

Figura 25. Validación de la información



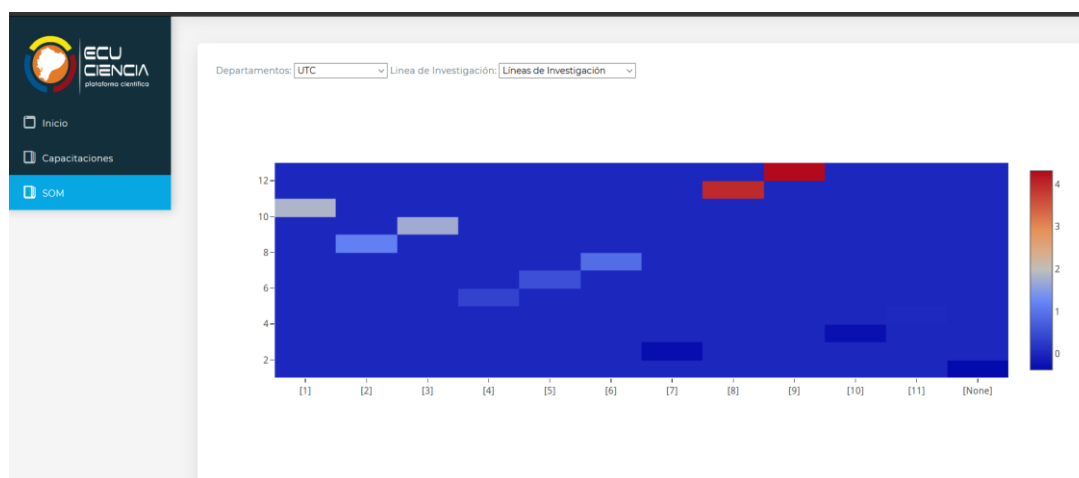
Elaborado por: Alex Geovanny Rivera Chasiquiza

La validación del entrenamiento y agrupación es importante para obtener resultados pulcros en la interfaz que interactúa directamente con los usuarios. Los datos procesados son convenientes sean validados para que la gráfica tenga mayor veracidad y sean fácilmente interpretados.

ETAPA 5. Visualización:

La visualización es también importante dentro del método de trabajo, ya que la investigación está enmarcada en interpretar los resultados. Es por esto importante que todas las fases mencionadas en este método de trabajo sean realizadas correctamente. Además, para el caso de la visualización, para una mejor interpretación de los resultados graficarlos mediante mapas de calor como se muestra en la Figura 28.

Figura 26. Visualización de información procesada



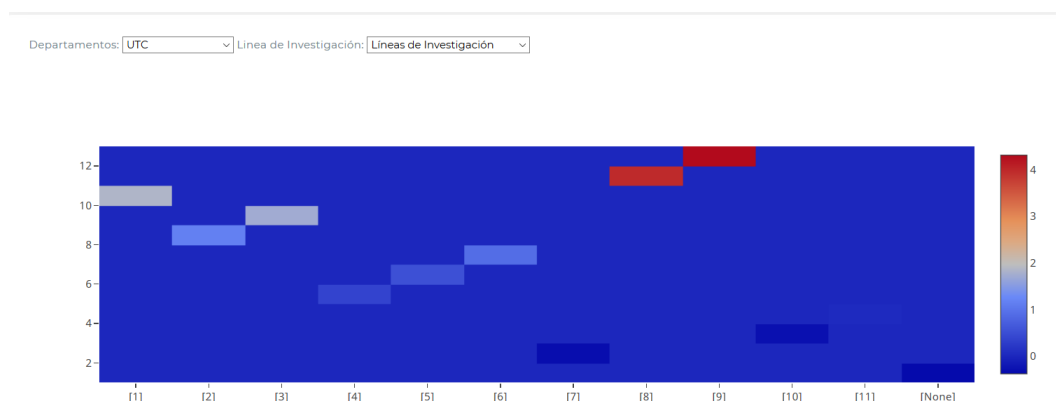
Elaborado por: Alex Geovanny Rivera Chasiquiza

En las figuras 29 y 30 expuestas a continuación se puede evidenciar la interfaz que interactúa directamente con el usuario. En las figuras 29 y 30 se muestra el resultado del método de trabajo alineado a SOM en el caso de las Líneas y Sublíneas de investigación según la Universidad Técnica de Cotopaxi correspondientemente, en las que se interpreta acorde a la paleta de colores ubicada al lado derecho del mapa de calor. En la figura 29 se puede evidenciar que de acuerdo al mapa de calor la línea de investigación en la que más artículos se encuentran registrados es la 9, la misma que corresponde a Educación, comunicación y diseño gráfico para el desarrollo humano y social. Y por su parte la Línea de investigación que menos es abordada por los investigadores en la Universidad Técnica de Cotopaxi es la 10, la cual corresponde a Cultura, patrimonio y saberes ancestrales.

En la figura 30 se evidencia las sublíneas de investigación, la cual se interpreta que la sublínea con mayor intervención de investigaciones en la Universidad Técnica

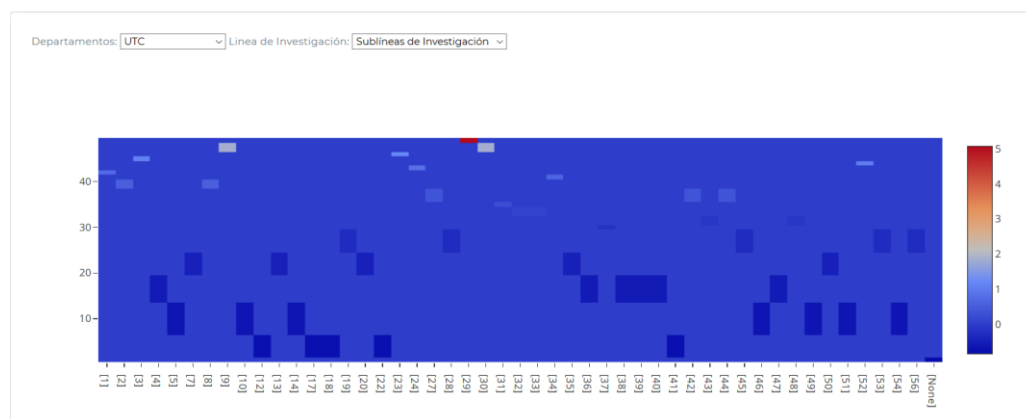
de Cotopaxi es la 29, la cual es correspondiente a Estrategias administrativas, productividad y emprendimiento. De igual manera se interpreta que las sublíneas a las que menos investigadores abordan son las 12 correspondiente a Investigación-innovación y emprendimientos (agroindustriales), 22 correspondiente a Turismo comunitario y 41 que corresponde a Comunicación organizacional.

Figura 27. Líneas de Investigación según la Universidad Técnica de Cotopaxi



Elaborado por: Alex Geovanny Rivera Chasiqiza

Figura 28. Sublíneas de Investigación según la Universidad Técnica de Cotopaxi



Elaborado por: Alex Geovanny Rivera Chasiqiza

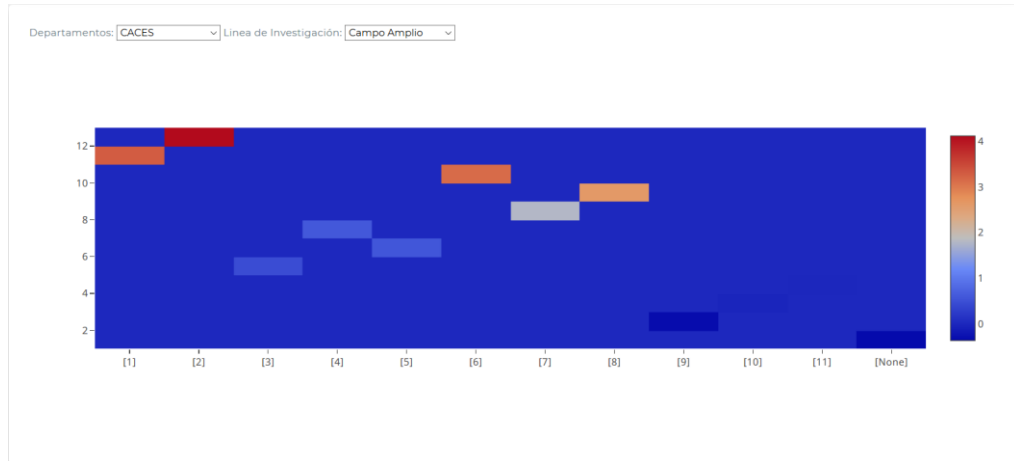
Por su parte en las figuras 31, 32 y 33 se identifican los campos amplios, específicos y detallados de acuerdo al CACES, en las cuales al igual que lo descrito anteriormente se visualizan mediante mapas de calor por medio del método de trabajo alineado a SOM. La figura 31 evidencia el resultado del Campo Amplio, en el cual se observa al código 2 correspondiente a Agricultura, silvicultura, pesca y

veterinaria como al que más investigadores se han destinado. De la misma manera se observa al código 9 correspondiente a Programas genéricos y calificaciones como el campo amplio que menos importancia le dan los investigadores.

En cuanto al campo específico se visualiza que los investigadores dirigen más su atención al código 1 correspondiente a Negocios y administración. Y de la misma manera a los que menos investigadores relacionan sus investigaciones, los cuales son el código 4 correspondiente a Silvicultura, código 8 correspondiente a Humanidades (excepto idioma) y código 20 correspondiente a Arquitectura y construcción.

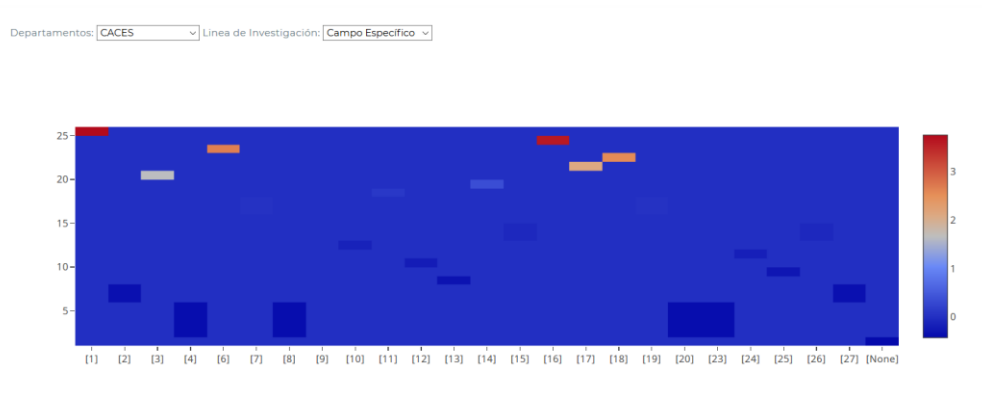
En los resultados visualizados referentes al campo específico se evidencia el código 13 correspondiente a Veterinaria como el que mayor número de investigadores tiene a sus servicios. De igual manera interpretando a los que menos investigadores llama la atención, encontramos al código 7 correspondiente a Habilidades laborales, código 11 correspondiente a Silvicultura, código 15 correspondiente a Bellas artes, código 31 correspondiente a Matemáticas, código 34 correspondiente a Ciencias políticas y educación cívica, código 55 correspondiente a Textiles (ropa, calzado y cuerpo), código 56 correspondiente a Minería y extracción, código 57 correspondiente a Arquitectura y urbanismo, código 58 correspondiente a Construcción e ingeniería civil, código 61 correspondiente a Desarrollo y habilidades personales, código 65 correspondiente a Tecnologías de diagnóstico y tratamiento médico, código 66 correspondiente a Terapia y rehabilitación y código 64 correspondiente a Enfermería y Obstetricia.

Figura 29. Campo Amplio según el CACES



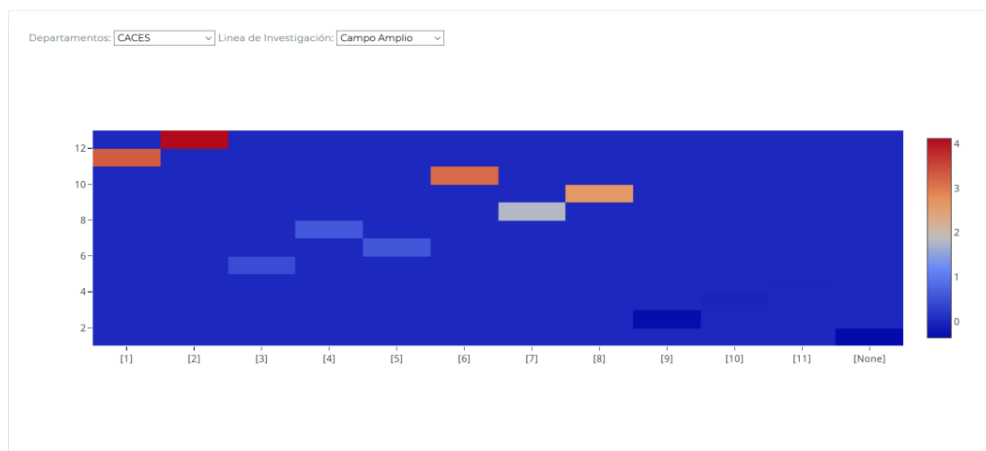
Elaborado por: Alex Geovanny Rivera Chasiquiza

Figura 30. Campo Específico según el CACES



Elaborado por: Alex Geovanny Rivera Chasiquiza

Figura 31. Campo Detallado según el CACES



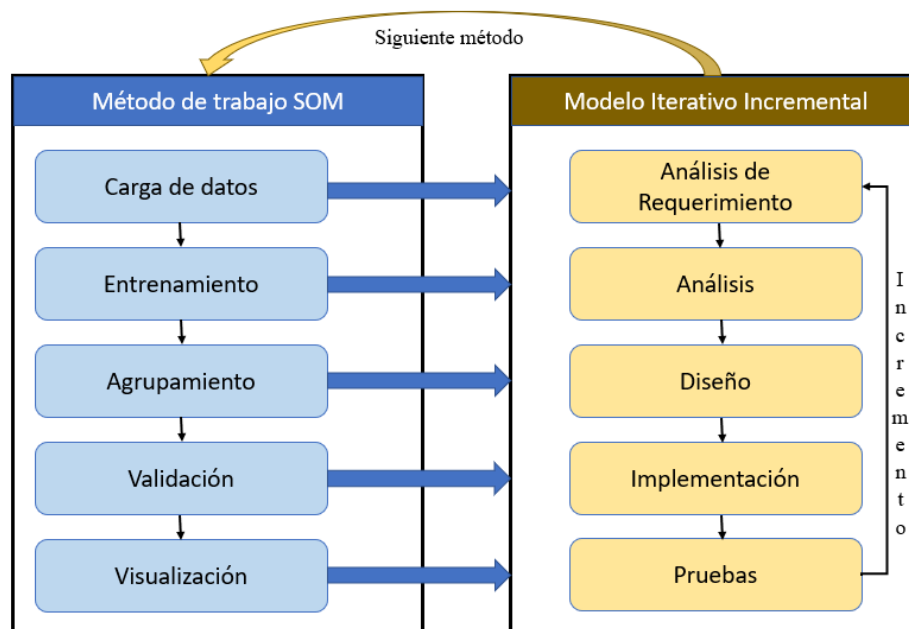
Elaborado por: Alex Geovanny Rivera Chasiquiza

Además, se mostrará aplicando este método de trabajo el nivel de impacto de los artículos científicos, ponencias y libros de acuerdo a la categorización que dicha revista pertenezca, siendo esto un valor agregado para la toma de decisiones debido a que mientras el investigador tenga indexado más artículos a revistas de alta calidad será premiado de acuerdo a las normas y estatutos de la Universidad Técnica de Cotopaxi.

El método de trabajo planteado es de gran ayuda para el proceso de investigación ya que permite interactuar de manera más amigable y eficaz con los requisitos obtenidos de las demás técnicas y con el desarrollo del módulo solicitado para el análisis de la data. Además, esta metodología queda abierta para ser utilizada por más investigaciones referentes a Inteligencia Artificial, Redes Neuronales Artificiales, Toma de Decisiones y Autoaprendizaje. La visualización de la data es recomendable se lo realice por mapas de calor, debido a su fácil interpretación entre el sistema y el usuario.

3.2.2. METODOLOGÍA DEL MODELO ITERATIVO INCREMENTAL

Figura 32. Etapas SOM con el modelo Iterativo Incremental



Elaborado por: Alex Geovanny Rivera Chasiqiza

A pesar que el modelo iterativo incremental es una metodología muy utilizada para el desarrollo de software, también se lo puede adaptar y utilizar para otros entornos. En la investigación se lo implemento en cada etapa del método de trabajo SOM (mapeo Auto-organizado), esto con el fin de mejorar en cierto modo la liberación de cada etapa planteada y así obtener resultados completamente funcionales a la salida hacía el usuario final, en este caso la visualización mediante los mapas de calor de la información procesada. Además, esta metodología permite realizar entregas de funcionales cada periodo de tiempo. Es así que a continuación se detallan las fases y sus respectivas implementaciones siguiendo estándares que ayudan a mejorar los procesos.

FASE I. Análisis de los requerimientos:

Para el análisis de requerimientos se utilizará la especificación de requerimientos de acuerdo a la plantilla IEEE, la cual favorece en este proceso siendo sumamente importante para la correcta recolección de información para el desarrollo correcto del módulo de visualización de los datos de producción científica de docentes investigadores en la Universidad Técnica de Cotopaxi.

Introducción:

En la Tabla 3 se especifican los requerimientos de software (ERS) para el módulo de visualización de datos de producción científica de docentes investigadores en la Universidad Técnica de Cotopaxi mediante mapeo auto-organizado, la misma que ha sido estructurada basándose en el estándar IEEE 830 el cual permite la correcta obtención de requerimientos para posteriormente continuar con la elaboración de diagramas útiles para el diseño y desarrollo del módulo.

Propósito:

La documentación presentada se encarga de definir los requisitos funcionales y no funcionales para el diseño e implementación del módulo de visualización de datos cuantitativos, tanto desde la entrada de la información por parte de los investigadores hasta la salida (visualización) de la misma por los usuarios que interactuaran con el sistema ECUCIENCIA.

Alcance:

La especificación de requerimientos está encaminada hacia los docentes investigadores y usuarios externos, por lo cual se toma como punto de partida para el desarrollo del módulo de visualización de datos de producción científica de docentes investigadores en la Universidad Técnica de Cotopaxi mediante mapeo auto-organizado, facilitando la interacción desde los investigadores hasta los usuarios externos.

Personal involucrado:*Tabla 4. Contenido de involucrados*

Nombre:	Alex Geovanny Rivera Chasiquiza
Rol:	Analista, diseñador y programador
Categoría Profesional:	<ul style="list-style-type: none">• Ingeniero en Informática y Sistemas Computacionales• Estudiante de Maestría de Sistemas de Información.
Responsabilidad:	Analista de requerimientos, diseño y programación del módulo de visualización de datos cuantitativos mediante SOM para el sistema web ECUCIENCIA.
Información de contacto:	alex.rivera2@utc.edu.ec

Elaborado por: Alex Geovanny Rivera Chasiquiza

Definiciones, acrónimos y abreviaturas:*Tabla 5. Definiciones, acrónimos y abreviaturas.*

NOMBRE	DESCRIPCIÓN
INVESTIGADOR	Docente investigador con libros, revistas científicas publicadas.
USUARIO EXTERNO	Persona que externamente a la Universidad Técnica de Cotopaxi visualizara la información subida por el investigador.
UTC	Universidad Técnica de Cotopaxi.

SOM	Mapeo auto-organizado.
ERS	Especificaciones de Requisitos de Software.
RF	Requisitos Funcionales.
RNF	Requisitos no Funcionales.
IA	Inteligencia Artificial.
RNA	Redes Neuronales Artificiales
BD	Almacenamiento de información.

Elaborado por: Alex Geovanny Rivera Chasiquiza

Referencias:

Tabla 6. Referencias

TÍTULO DEL DOCUMENTO	REFERENCIA
Standard IEEE 830-1998	IEEE

Elaborado por: Alex Geovanny Rivera Chasiquiza

Resumen:

El presente documento está estructurado en tres partes, la primera está relacionada hacia la introducción referente a lo que se propone realizar proporcionando una visión general de las especificaciones módulo tomando los requerimientos solicitados por los usuarios encargados del sistema ECUIENCIA. La segunda está destinada hacia la descripción general del sistema, esto para obtener las funcionalidades que el módulo debe realizar, de acuerdo a los requisitos presentados, datos asociados, restricciones, supuestos y dependencias para el desarrollo del mismo. Y por último la tercera, la cual está destinada hacia la definición detallada de los requisitos para satisfacer las necesidades de los usuarios, esto se realizará por medio de los requisitos funcionales y no funcionales.

Descripción general

Perspectiva del producto:

El módulo de visualización de datos de producción científica de docentes investigadores en la Universidad Técnica de Cotopaxi mediante mapeo auto-organizados, será diseñado para complementar el sistema web ECUCIENCIA, aportando mayor eficacia del mismo ya que permitirá mayor interacción entre el sistema y los usuarios que interactuaran directamente con el mismo. Esto permitirá un mejor posicionamiento con sistemas que ofrecen funcionalidades similares ya que la mayoría de estos no incluyen IA y tampoco cuentan con visualización de información mediante SOM.

Funcionalidad del producto:

Las funcionalidades del módulo propuesto para complementar el sistema web ECUCIENCIA es detallado a continuación mediante la presentación de los diagramas de caso de uso que nos facilita el entendimiento de las funcionalidades que se propone implementar.

Características de los usuarios:

Tabla 7. Usuario Investigador.

Tipo de usuario	Investigador
Formación	Docente Investigador.
Actividades	Encargado de subir sus libros, revistas y artículos científicos

Elaborado por: Alex Geovanny Rivera Chasiqiza

Tabla 8. Usuario

Tipo de usuario	Usuario
Formación	Usuario externo.
Actividades	Visualizar la información presentada por el sistema ECUCIENCIA.

Elaborado por: Alex Geovanny Rivera Chasiqiza

Restricciones:

- Interfaz web.
- Autenticación en el sistema.
- Módulo para ser usado con internet.
- Lenguajes y tecnologías de uso: HTML, Python, Django y PostgreSQL.
- Base de Datos preparada para la cantidad de transacciones que sea necesaria.
- Sistema basado en MVC.
- Sistema diseñado e implementado con ingeniería de usabilidad.

Suposiciones y dependencias:

- Requisitos descritos estables y óptimos para la trazabilidad entre los usuarios y el sistema.
- Servidor a implementar debe cumplir la especificación de requerimientos mínimos para su uso.
- Conexión estable a internet.

Requisitos específicos

Requerimientos Funcionales:

Tabla 9. RF01

Identificación del requerimiento del RF01	
Nombre del requerimiento:	Ingreso al Sistema
Características:	El investigador debe autenticarse para acceder al sistema.
Descripción del requerimiento:	Los datos de producción científica podrán ser ingresados únicamente por los investigadores.
Requerimiento no funcional	
Prioridad del requerimiento	Alta

Elaborado por: Alex Geovanny Rivera Chasiquiza

Tabla 10. RF02

Identificación del RF02 requerimiento		
Nombre del requerimiento:	del	Procesar información de UTC
Características:		El sistema gestiona la información basada en el método de trabajo alineado a SOM de la Universidad Técnica de Cotopaxi.
Descripción del requerimiento:	del	EL sistema se encarga de visualizar el mapa de calor de las líneas y sublíneas de investigación.
Requerimiento funcional	no	
Prioridad del requerimiento	del	Alta

Elaborado por: Alex Geovanny Rivera Chasiquiza

Tabla 11. RF03

Identificación del RF03 requerimiento		
Nombre del requerimiento:	del	Procesar información de CACES
Características:		El sistema gestiona la información basada en el método de trabajo alineado a SOM del CACES.
Descripción del requerimiento:	del	EL sistema se encarga de visualizar el mapa de calor del campo amplio, específico y detallado según el CACES.
Requerimiento funcional	no	
Prioridad del requerimiento	del	Alta

Elaborado por: Alex Geovanny Rivera Chasiquiza

Tabla 12. RF04

Identificación del requerimiento		del RF04
Nombre del requerimiento:		Procesar información de UNESCO
Características:		El sistema gestiona la información basada en el método de trabajo alineado a SOM de la UNESCO
Descripción del requerimiento:		EL sistema se encarga de visualizar el mapa de calor del campo amplio, específico y detallado según el UNESCO.
Requerimiento funcional	no	
Prioridad del requerimiento		Alta

Elaborado por: Alex Geovanny Rivera Chasiquiza

Tabla 13. RF05

Identificación del requerimiento		del RF05
Nombre del requerimiento:		Procesar información del nivel de impacto.
Características:		El sistema gestiona la información basada en el método de trabajo alineado a SOM del nivel de impacto de los artículos.
Descripción del requerimiento:		EL sistema se encarga de visualizar el mapa de calor el nivel de impacto de los artículos según la categorización de la revista en la que fue subido, esto depende directamente del investigador.
Requerimiento funcional	no	
Prioridad del requerimiento		Alta

Elaborado por: Alex Geovanny Rivera Chasiquiza

Requerimientos no Funcionales:

Tabla 14. RNF01

Identificación del requerimiento		RNF01
Nombre del requerimiento:	del	Interfaz del sistema
Características:		El sistema presenta interfaces sencillas y comprensibles para el manejo del usuario.
Descripción del requerimiento:	del	Para que la interacción entre el sistema y el usuario sea efectiva las interfaces propuestas son sencillas de manejarlas, además, las interfaces son responsive permitiendo el acoplamiento de acuerdo a la resolución de pantalla.
Prioridad del requerimiento	del	Alta

Elaborado por: Alex Geovanny Rivera Chasiqiza

Tabla 15. RNF02

Identificación del requerimiento		RNF02
Nombre del requerimiento:	del	Mantenimiento
Características:		El sistema contara con la documentación incluida en las líneas de código.
Descripción del requerimiento:	del	El sistema contara con la documentación incluida en las líneas de código para el mantenimiento correcto en caso de necesitarlo.
Prioridad del requerimiento	del	Alta

Elaborado por: Alex Geovanny Rivera Chasiqiza

Tabla 16. RNF03

Identificación del requerimiento		RNF03
----------------------------------	--	--------------

Nombre del requerimiento:	Desempeño
Características:	El sistema será eficaz en el procesamiento de los datos.
Descripción del requerimiento:	El sistema contara con un desempeño fiable para el usuario con el que interactúa. La información estará siempre disponible y actualizada para la visualización.
Prioridad del requerimiento	Alta

Elaborado por: Alex Geovanny Rivera Chasiquiza

Tabla 17. RNF04

Identificación del requerimiento del RNF04	
Nombre del requerimiento:	Usuarios
Características:	El sistema se basará de acuerdo al usuario ingresado.
Descripción del requerimiento:	El sistema garantiza la información sea visualizada de acuerdo al usuario ingresado, además, garantiza el nivel del mismo (Administrador o Usuario).
Prioridad del requerimiento	Alta

Elaborado por: Alex Geovanny Rivera Chasiquiza

Tabla 18. Confiabilidad

Identificación del requerimiento del RNF05	
Nombre del requerimiento:	Confiabilidad
Características:	El sistema será plenamente confiable en sus funcionalidades.
Descripción del requerimiento:	El sistema garantiza el continuo funcionamiento sin interrupciones, salvo casos especiales en el que se tenga que dar m mantenimiento al servidor en el que reposa el mismo.

Prioridad del requerimiento	Alta
-----------------------------	------

Elaborado por: Alex Geovanny Rivera Chasiquiza

Tabla 19. Seguridad

Identificación del requerimiento		RNF06
Nombre del requerimiento:	del	Seguridad
Características:		El sistema proporciona seguridad en los datos procesados.
Descripción del requerimiento:	del	El sistema avala la seguridad de los datos almacenados en su gestor de base de datos.
Prioridad del requerimiento	del	Alta

Elaborado por: Alex Geovanny Rivera Chasiquiza

Requisitos comunes de las interfaces

Interfaces de usuario:

La interfaz de usuario pretende ser completamente funcional y amigable con el usuario. Para una mejor interacción con el usuario las interfaces constarán de pestañas que faciliten la movilización por la misma, además, para los filtros se implementarán combobox que permitan seleccionar lo necesario, haciendo que sea completamente manejable.

Interfaces de Hardware:

Para el funcionamiento correcto será necesario navegar en el sistema ECUCIENCIA desde un computador o dispositivos inteligentes. EN cuanto al computador es recomendable tener los siguientes requisitos:

- Procesador superior a 1.66 GHz.
- Memoria RAM superior a 500 Mb. (para su mejor desempeño).
- Acceso a Internet (el internet debe ser estable).
- Mouse

- Teclado

Para el caso de otro dispositivo inteligente será necesario que tenga los siguientes requerimientos:

- Espacio de memoria para la navegación.
- Conexión a Internet (la conexión debe ser estable)

Interfaces de Software:

El software a utilizarse al ser un sistema web no muestra demasiada importancia el software que tenga instalado, sin embargo, se recomienda tener el software actualizado y desinfectado (con el antivirus de su confianza).

Interfaces de comunicación:

La comunicación se los realiza mediante protocolos de internet estándares para su manejo. Por lo que es importante la conexión a internet.

FASE II. Análisis

Esta fase se encarga del análisis de los requerimientos planteados en la fase anterior, a partir de los cuales se identifican las historias de usuarios y diagramas de caso de uso, los mismos que serán de mucha ayuda en el desarrollo de la solución para la visualización de datos de producción científica referentes a los docentes investigadores en la Universidad Técnica de Cotopaxi, el mismo que estará incluido en el sistema ECUCIENCIA.

Historia de Usuarios

Las historias de usuario surgen a partir de la entrevista realizada al PhD. Gustavo Rodríguez y al Ing. Alex Cevallos como responsables del sistema ECUCIENCIA. En el Anexo II encontraremos la Tabla 21 y la Tabla 22, las cuales muestran las principales historias de usuario para el módulo de visualización mediante SOM que será integrado al sistema ECUCIENCIA.

Diagrama de casos de uso

En los Anexo III podemos visualizar la Figura 23 con el diagrama general de casos de uso de acuerdo a los requerimientos planteados y analizados previamente.

FASE III. Diseño

En el diseño se lo realiza a partir de tener establecidas las iteraciones. En esta fase se desarrollan los diagramas de clases, secuencias y actividades, dependiendo el caso de ser necesarias. En el caso para el desarrollo del módulo de visualización mediante SOM de los datos de producción científica en la Universidad Técnica de Cotopaxi se implementarán solo los diagramas de clases y secuencia. En el Anexo IV se puede apreciar los diagramas mencionados en la Figura 24 y la Figura 25 respectivamente.

FASE IV. Implementación

Esta fase es igual de importante que las mencionadas, debido a que es aquí donde se pone en práctica las fases anteriores para desarrollar en un lenguaje de programación establecido y conectado a una base de datos. Se programa la solución a la problemática planteada en la investigación. En el caso del módulo a desarrollarse para la visualización mediante mapeo auto-organizado de datos de producción científica en la Universidad Técnica de Cotopaxi, en el sistema ECUCIENCIA será desarrollada en Python como se mencionó con el framework Django y como gestor de BD PostgreSQL. En el Anexo V se pueden apreciar la Figura 26, Figura 27, Figura 28 y Figura 29 con los resultados obtenidos en el desarrollo del módulo.

FASE V. Pruebas

En esta fase se valida los resultados acordes a los requisitos tomados en las fases anteriores. Para esta validación es necesario utilizar una matriz de pruebas que se las puede observar en el Anexo VI en la Tabla 23 y Tabla 24 respectivamente. Estas fichas las realiza el tester.

3.3. RESULTADOS DEL DISEÑO EXPERIMENTAL

3.3.1. MÉTODO DE CRITERIO DE EXPERTOS

Método Delphi

La validación de la propuesta tecnológica se la realizo utilizando el método Delphi debido a que es importante asegurarse que la investigación cumpla las expectativas planteadas. Por lo cual al desarrollar las fases indicadas en la metodología es prudente aplicar instrumentos de validación, motivo por el cual investigadores

recomiendan calcular el estado de fiabilidad a los expertos, y resultados de los cuestionarios que se les aplica y así reflejar resultados fiables para su puesta en producción. Estas encuestas tuvieron un enfoque cualitativo ya que la información recolectada no cuenta con mediciones numéricas sino más bien los expertos respondieron de acuerdo a los atributos y características de la parte teórica y práctica de la investigación. El método Delphi está compuesto de cinco fases que serán expresadas a continuación.

Fase preliminar

Permite separar por grupos a los responsables de validar la investigación, en la cual se dividió en dos grupos. El primer grupo está compuesto por el encargado de poner en marcha la propuesta y seleccionara a los expertos para validar los resultados. El segundo grupo está conformado por los expertos que a través de su experiencia verifican y expresan sus conocimientos en el área de estudio mediante los cuestionarios que se los hace llegar para que sean llenados.

Todos los procesos se los realizo mediante sus respectivos correos electrónicos, en el cual se le adjuntó un link con las encuestas correspondientes subidas a hojas de cálculo que la herramienta Google Drive nos ofrece, en dicho documento en la primera página se brindó una breve introducción al tema de investigación, problemática, objetivos, métodos de trabajo y un efímero concepto acerca de los Mapas Auto-organizados (SOM). La siguiente página estuvo compuesta por el cuestionario correspondiente para obtener el coeficiente de competencia del experto. Las siguientes páginas están destinadas para los cuestionarios que permiten validar la propuesta práctica y teórica de la investigación planteada. Cabe recalcar que las encuestas fueron totalmente personales y anónimas para evitar conflictos y obtener mejores resultados.

Selección de expertos

Esta fase es primordial ya que de la correcta selección de los expertos permite obtener mejores resultados, debido a esto se tomó en cuenta la profesión, cargo y años de experiencia de cada experto en sus respectivas áreas. El grupo de expertos quedo conformado por un magister en docencia universitaria y administración

educativa, un magister en gerencia en sistemas de información, dos ingenieros en informática y sistemas computacionales, dos ingenieros en sistemas y un ingeniero en software. En total se seleccionaron a siete expertos, cada uno de los cuales poseen cargos relevantes en sus áreas. Desde docentes académicos, gerentes de empresas importantes, jefes en sus respectivas áreas hasta desarrolladores de software. Ver tablas 19 y 20.

Tabla 20. Resultados del procesamiento para la determinación del coeficiente de competencia de expertos

Expertos	Kc	Ka	K	Valoración		
1	0,9	0,92	0,91	Promedio de índice de competencia de expertos = 0,76 Kc = Medio	Kc	Coeficiente de Conocimiento
2	0,9	0,83	0,87		Ka	Coeficiente de Argumentación
3	0,9	0,83	0,87			
4	0,9	0,58	0,74		K	Coeficiente de Competencia K = 0,5(Kc + Ka)
5	0,8	0,67	0,73			
6	0,7	0,75	0,73			
7	0,8	0,58	0,69			
Si $0,8 < K < 1,0$ Coeficiente de competencia alto						
Si $0,5 < K < 0,8$ Coeficiente de competencia medio						
Si $K < 0,5$ Coeficiente de competencia bajo						

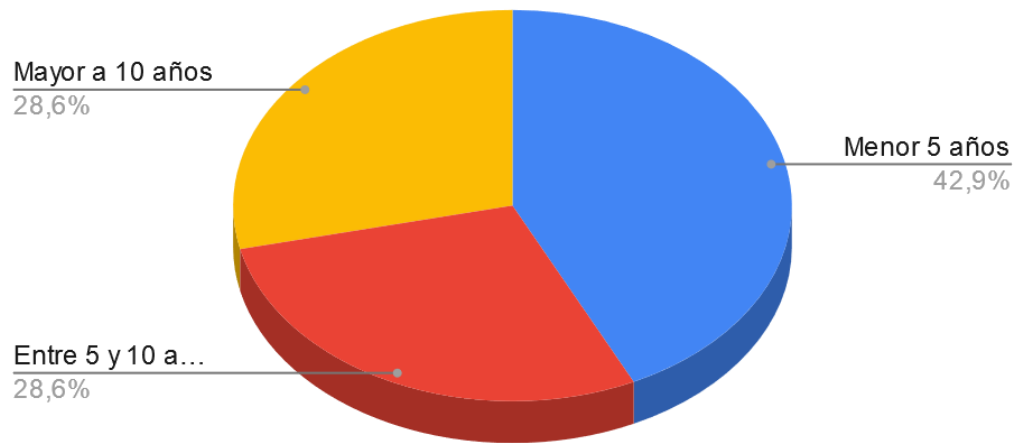
Elaborado por: Alex Geovanny Rivera Chasiqiza

Tabla 21. Caracterización de los expertos que participaron en la consulta

N°	Calificación personal	Categoría docente y/o cargo	Años de experiencia
1	Magister Docencia Universitaria y Administración Educativa	Docente	16
2	Ingeniero en Informática y Sistemas Computacionales	Desarrollador de Software	2
3	Ingeniero en Informática y Sistemas Computacionales	Desarrollador de Software	2
4	Ingeniero en Sistemas	Gerente	27
5	Ingeniero en Sistemas	Supervisor Área Técnica	9
6	Ingeniero en Software	Desarrollador de Software	3
7	Magister en Gerencia en Sistemas de Información	Jefe Área Desarrollo de Software	8

Elaborado por: Alex Geovanny Rivera Chasiquiza

Figura 33. Años de experiencia en su respectiva área



Elaborado por: Alex Geovanny Rivera Chasiquiza

Los expertos seleccionados para la validación de la investigación provienen de distintas instituciones tanto públicas como privadas, de los cuales un experto es un docente académico como se puede observar en la tabla 20 y los demás expertos están relacionados directamente con el desarrollo de software y manejo de grandes cantidades de datos. El índice de competencia referente a la propuesta de investigación es medio como se puede apreciar en la tabla 21. El rango para el proceso estadístico de la información recolectada con el método Delphi fue entre 0 y 1, tabulando los resultados en la tabla 22.

Fase exploratoria

Los cuestionarios a aplicarse fueron desarrollados tomando en cuenta los aspectos críticos en la investigación tanto práctica como teórica. En la primera encuesta se solicitó a los expertos en un rango establecido dependiendo la respectiva pregunta seleccionara sus conocimientos en torno a la temática de investigación planteada referente al manejo de grandes cantidades de datos en la actualidad a nivel personal y sus criterios respecto a investigaciones desarrolladas o en desarrollo a nivel

nacional e internacional, estos resultados sirvieron para evidenciar el nivel de conocimiento respecto al tema de cada experto.

En el segundo cuestionario se planteó preguntas referentes al mapeo auto-organizado y sus criterios a cerca de ponerlos en producción en la Universidad Técnica de Cotopaxi para clasificar y visualizar los datos de producción científica según sus líneas y sublíneas de investigación, además de los campos amplios, específicos y detallados de acuerdo al CACES y UNESCO. También, se vio prudente realizar preguntas referentes al módulo SOM para la visualización de esta información en el sistema ECUCIENCIA que actualmente se encuentra en producción en la Universidad Técnica de Cotopaxi. Ver tabla 21.

Fase final

En esta fase se sintetizaron los resultados del proceso de validación con el cuestionario final expuesto a los expertos. El cuestionario definitivo se lo aprecia en la tabla 21, con el cual se obtuvo la valoración correspondiente de la investigación planteada acerca de la visualización de datos de producción científica de docentes investigadores en la Universidad Técnica de Cotopaxi según sus líneas y sublíneas de investigación, y de acuerdo a los campos amplios, específicos y detallados de acuerdo a la UNESCO y CACES. Se recalca que para verificar la fiabilidad del cuestionario fue necesario utilizar metodologías que estudiosos proponen para la correcta aplicación del método Delphi.

Resultados

Los resultados obtenidos son expresados a continuación mediante la tabulación y validación de los cuestionarios planteados desde la elección de expertos según su coeficiente de competencia, caracterización de los expertos participantes, años de experiencia y cuestionario con los criterios definitivos para evaluar la propuesta tecnológica acerca de la visualización de datos de producción científica de docentes investigadores en la Universidad Técnica de Cotopaxi según sus líneas y sublíneas de investigación, y los campos amplios, específicos y detallados de acuerdo a la UNESCO y CACES.

Tabla 22. Definición de los criterios y escala para evaluar el método SOM.

Código	Preguntas	Escala de Evaluación				
		R1	R2	R3	R4	R5
R1: No Adecuado R2: Poco Adecuado R3: Adecuado R4: Bastante Adecuado R5: Muy Adecuado						
MÉTODO DE MAPEO AUTO-ORGANIZADO (SOM)						
En vista a la introducción de Inteligencia Artificial en el manejo de datos en la actualidad.						
PRE1	¿Cree importante el análisis, clasificación y visualización de grandes cantidades de información?					
PRE2	¿Piensa se debería clasificar la información de docentes investigadores en la Universidad Técnica de Cotopaxi de acuerdo a los campos de investigación de la UNESCO, CACES y Universidad Técnica de Cotopaxi respectivamente?					
PRE3	¿Le parece que se debería utilizar métodos para realizar los procesos de análisis, clasificación y visualización de información de los campos mencionados?					
PRE4	¿El método de Mapeo Auto-organizado (SOM) sería adecuado para estos procesos?					
MÓDULO SOM PARA SISTEMA ECUCIENCIA						
Visualización de información mediante Mapeo Auto-organizados (SOM)						
FUNC	¿Usted está de acuerdo con el funcionamiento del Módulo propuesto con visualización mediante SOM?					
FIAB	¿Usted considera el Módulo tiene un rendimiento adecuado para el manejo de grandes cantidades de datos?					
USAB	¿Considera el Módulo es atractivo, intuitivo y de fácil uso para el usuario?					
EFIC	¿Usted está de acuerdo que el rendimiento del Módulo propuesto relativo a la cantidad de recursos utilizados es adecuado?					
MANT	¿Considera el Módulo propuesto será fácilmente modificado y adaptado al sistema ECUCIENCIA en el futuro?					

PORT	¿Usted estima Módulo podrá ser transferido a otro ambiente en el caso de ser necesario?									
------	---	--	--	--	--	--	--	--	--	--

Elaborado por: Alex Geovanny Rivera Chasiquizza

Tabla 23. Estadísticos descriptivos de la valoración del cuestionario de evaluación.

EXPERTO	PRE1	PRE2	PRE3	PRE4	FUNC	FIAB	USAB	EFIC	MANT	PORT	SUMA
1	5	5	5	5	5	4	4	4	4	4	45
2	5	4	4	4	5	4	5	4	5	5	45
3	5	5	5	5	5	4	4	5	4	5	47
4	5	3	4	4	4	4	4	4	4	4	40
5	5	4	5	5	5	5	5	5	5	5	49
6	5	4	5	4	4	5	5	5	5	4	46
7	5	5	5	5	5	4	5	4	5	5	48
VARIANZAS	0,00	0,49	0,20	0,24	0,20	0,20	0,24	0,24	0,24	0,24	

Elaborado por: Alex Geovanny Rivera Chasiquizza

Para la validación de las respuestas marcadas por los expertos es importante tomar en cuenta que las puntuaciones colocadas en el cuestionario no son suficiente evidencias de la validez que este tenga, por tal motivo ciertos investigadores del método Delphi aconsejan determinar el coeficiente de Spearman-Brown y el alfa de Cronbach. En la Tabla 23 se puede apreciar dicho cálculo del coeficiente.

Tabla 24. Alfa de Cronbach

Resultados	
α (Alfa) =	0,76
K (número de items) =	10
V_i (varianza Inicial) =	2,33
V_t (varianza total) =	7,35

Elaborado por: Alex Geovanny Rivera Chasiquizza

La fórmula ocupada para calcular el Alfa de Cronbach es la siguiente propuesta por los conocedores del método Delphi:

$$\alpha = \frac{K}{K - 1} \left[1 - \frac{\sum V_i}{V_t} \right]$$

Los resultados del desarrollo del módulo para la visualización de datos de producción científica acorde a los docentes investigadores en la Universidad Técnica de Cotopaxi se los puede evidenciar en la interfaz gráfica integrada al sistema ECUCIENCIA. Cabe mencionar que, al haber utilizado una metodología para el desarrollo de la propuesta, es prudente analizar los resultados obtenidos en dicha metodología.

Al analizar los resultados obtenidos en el Método de trabajo alineado al Mapeo auto-organizado (SOM), el cual fue diseñado específicamente para esta investigación se puede evidenciar que es completamente fiable al integrarlo en una investigación relacionada a mapeo autoorganizado, redes neuronales artificiales, toma de decisiones y autoaprendizaje. Esto se debe a que está orientada al manejo de grandes cantidades de datos complementándola con el análisis, entrenamiento y visualización de los mismos, esto conlleva a interpretar que el método de trabajo alineado a SOM propuesta se podrá utilizar en posteriores investigaciones más a fondo en los temas planteados como principales referentes.

3.4. RESULTADOS DE LA VALORIZACIÓN ECONÓMICA, TECNOLÓGICA, OPERACIONAL Y AMBIENTAL

Para determinar la factibilidad de implementar la metodología propuesta del mapeo auto-organizado (SOM) en el desarrollo del módulo de visualización de datos de producción científica en la Universidad Técnica de Cotopaxi es importante considerar las valorizaciones invertidas en la investigación de la propuesta, las mismas son detalladas a continuación:

3.4.1. VALORACIÓN ECONÓMICA:

Lo económico es importante en el desarrollo de una investigación, en especial para el desarrollo de un software, en este sentido a continuación se detallan los gastos directos e indirectos.

Gastos Directos:

Tabla 25. Gastos directos

Descripción	Gastos	Detalle	Cantidad	V. Unitario	Total
Software	Software	Python	1	Licencia gratuita	\$ 0
		PostgreSQL	1	Licencia gratuita	\$ 0
		Navegador de Internet	1	Licencia gratuita	\$ 0
		Paquete de Office 2019 (Documentación)	1	\$ 163.30	\$ 163.30
		Visual Paradigm Community (Diagramación)	1	Licencia gratuita	\$ 0
		Internet	12 meses	\$ 20	\$ 240
Sistema	Módulo	Desarrollo	4 funciones	-----	\$ 4000
Papelería	Impresiones		450	\$ 0.05	\$ 22.5
	Copias		180	\$ 0.02	\$ 3.60
	Carpetas		5	\$ 0.75	\$ 3.75
	Sobres de manila		5	\$ 0.50	\$ 2.50
	Anillados		5	\$ 0.80	\$ 4.00
Servidor	Procesador	1 GHz			\$ 300

	Memoria	512 RAM	
	Espacio disco duro	8 Gb	
	Unidad	USB	
	Periféricos de entrada, salida y almacenamiento	Resolución 800 x 600 como mínimo	
Total:			\$ 4,739.65

Elaborado por: Alex Geovanny Rivera Chasiqiza

Los gastos directos empleados en la implementación del método de trabajo alineado a SOM y el módulo de visualización de datos de producción científicos en la Universidad Técnica de Cotopaxi ascienden a \$ 4,739.65, divididos de acuerdo a lo requerido durante el desarrollo de la investigación, en los cuales incluyen los softwares utilizados, costo de desarrollo, papelería y servidor. El desarrollo del módulo es el que a más valor asciende debido a las horas utilizadas en la investigación y programación para tener un resultado completamente funcional.

Gastos Indirectos:

Los gastos indirectos requeridos durante el proceso de investigación de la propuesta son detallados a continuación:

Tabla 26. Gastos Indirectos

Descripción	Total
Movilidad	\$ 300
Alimentación	\$ 50
Total:	\$ 350

Elaborado por: Alex Geovanny Rivera Chasiqiza

Gastos Totales:

Los gastos totales son descritos en la Tabla 27 a continuación:

Tabla 27. Gastos Totales

Descripción	Total
Gastos Directos	\$ 4,739.65
Gastos Indirectos	\$ 350
Imprevistos (10%)	\$ 508.97
Total:	\$ 5598.62

Elaborado por: Alex Geovanny Rivera Chasiqiza

3.4.2. VALORACIÓN TECNOLÓGICA:

En la actualidad la tecnología se ha convertido en parte fundamental en la vida estudiantil como profesional, por lo que el aporte que esta hace a los procesos de investigación es innegable. Para solucionar la problemática planteada en la propuesta de investigación respecto a la visualización de datos de producción científica en la Universidad Técnica de Cotopaxi, se toma en cuenta el impacto que tendrá en los usuarios que interactuaran directa o indirectamente con el módulo del sistema ECUCIENCIA. Este impacto tecnológico será principalmente visible por la utilización del método de trabajo alineado a SOM para el análisis, aprendizaje y visualización de la información mediante mapas de calor.

3.4.3. VALORACIÓN AMBIENTAL:

En cuanto a la valoración ambiental y el impacto que está conlleva, se puede mencionar como punto fundamental la disminución de materiales físicos como hojas impresas, plásticos, esferos, entre otros. Con lo expresado es visible este impacto, adicional a esto se tendrá presente los artículos subidos al sistema ECUCIENCIA en cualquier lugar que tenga acceso a Internet facilitando enormemente el traslado de información.

3.5. DISCUSIÓN DE LA APLICACIÓN Y VALIDACIÓN DE LA PROPUESTA

La investigación planteada al inicio de la propuesta buscaba implementar funcionalidades acorde tecnologías que actualmente están siendo investigadas y

puestas en práctica en todas partes del mundo, es así que se consideró al mapeo auto-organizado como una funcionalidad que brindaría al sistema ECUCIENCIA una ventaja considerable para graficar y analizar la información de los docentes investigadores en la Universidad Técnica de Cotopaxi. A partir de esta necesidad se planteó el desarrollo del módulo de visualización de datos de producción científica mediante mapeos auto-organizados tomando como punto de partida las líneas y sublíneas de investigación en la Universidad Técnica de Cotopaxi, y los campos amplios, específicos y detallados de acuerdo a la CACES y UNESCO respectivamente.

En el transcurso de la investigación se evidencio la carencia de una metodología que permita analizar el proceso de mapeo auto-organizado, por lo que se planteó el diseño y desarrollo de una metodología que acoja el concepto del aprendizaje no supervisado y toma de decisiones mediante la utilización de redes neuronales. A partir de esta investigación fue concebida el método de trabajo alineado al mapeo auto-organizado (SOM) con la cual se plantea desde el análisis y clasificación previo a la carga de datos en la librería a utilizar para el entrenamiento de la red neuronal, continuando con el agrupamiento de este aprendizaje y validando los resultados obtenidos para visualizarlos mediante mapas de calor que sean de fácil comprensión y análisis para el usuario que será el que interactuara directamente con el sistema ECUCIENCIA y con el módulo en cuestión.

Cabe recalcar que por cada etapa del método de trabajo planteado se implementó el modelo iterativo incremental, esto con el fin de mejorar los resultados antes de continuar con la siguiente etapa. Al desarrollar este método de trabajo enfocado principalmente al análisis e implementación de grandes cantidades de datos se dio solución a falencias detectadas en investigaciones realizadas por diferentes expertos en distintas partes del planeta, tales como la veracidad de la respuesta obtenida en la visualización de los mapas de calor, adaptación del método SOM a la Web, entre otros, lo que permite adecuar a la tecnología actual y dar solución a propuestas similares de investigadores en el futuro.

La validación de la propuesta fue el desarrollo del módulo que al estar implementado con el método de trabajo alineado a SOM mencionada es completamente funcional y eficaz, ya que muestra los datos en tiempo real mediante

mapas de calor. Esto demuestra que el método de trabajo alineado a SOM brindará gran ayuda en investigaciones futuras y mejorará el sistema ECUCIENCIA en cuanto a toma de decisiones automáticas y aprendizaje no supervisado se refiere. Con respecto al módulo se pudo probar el manejo de grandes cantidades de datos con respecto a las líneas y sublíneas de investigación en la Universidad Técnica de Cotopaxi y los campos amplios, específicos y detallados según el CACES y UNESCO respectivamente, con el valor adicional que corresponde al impacto de artículos de acuerdo a la categorización de revista a la que fue indexado.

3.6. CONCLUSIONES CAPITULO III

- Las metodologías planteadas tanto para el desarrollo del módulo de visualización como el método de trabajo alineado a SOM para el desarrollo de la investigación fueron importantes para resolver el problema planteado.
- El método de trabajo alineado al mapeo auto-organizado (SOM) como se esperaba brindo la ayuda necesaria para la investigación planteada, debido a su acoplamiento desde el análisis de la data, entrenamiento de la red neuronal y visualización de los resultados mediante mapas de calor.
- La visualización de la información mediante mapas de calor siguiendo el esquema del método de trabajo alineado a SOM fueron interactivos y eficaces.

CONCLUSIONES GENERALES

- El análisis de los resultados expuestos por diferentes investigadores a nivel mundial referente a redes neuronales artificiales y al mapeo auto-organizado fue beneficioso para conocer los antecedentes en cuanto al tema existían, esta información permitió generar propuestas para el desarrollo de la investigación.
- El diseño y desarrollo del método de trabajo alineado al mapeo auto-organizado (SOM) permitió el progreso de la solución con mayor fluidez, ya que interactúa con el usuario desde el ingreso de la información al sistema ECUCIENCIA hasta la visualización de la misma ya procesada y entrenada por medio de redes neuronales artificiales de acuerdo a las líneas y sublíneas de investigación en la Universidad Técnica de Cotopaxi y los campos amplios, específicos y detallados según la CACES y UNESCO.
- Los resultados obtenidos con la utilización del método de trabajo alineado al mapeo auto-organizado (SOM) son visualizados mediante mapas de calor en la interfaz del sistema ECUCIENCIA, en el módulo desarrollado como solución a la problemática con datos de producción científica según las líneas de investigación para la Universidad Técnica de Cotopaxi y los campos de acuerdo a la CACES y UNESCO.

RECOMENDACIONES

- Es importante tener en cuenta los antecedentes respecto al mapeo auto-organizado y las metodologías utilizadas para dar solución a las problemáticas.
- En la actualidad las redes neuronales artificiales están abriéndose campo en los sistemas web principalmente, es por este motivo conveniente la utilización de metodologías acorde a la problemática que se pretende solucionar.
- Los mapas auto-organizados (SOM) permiten interpretar grandes cantidades de datos con resultados más cercanos a la realidad, por lo que es prudente que los sistemas web lo implementen para tener mejor interacción con los usuarios.

BIBLIOGRAFÍA

- [1] L. Yuan, «Implementation of Self-Organizing Maps with Python,» *DigitalCommons@URI*, vol. 1, p. 124, 2018.
- [2] Sevamoo, «GITHUB,» 19 Septiembre 2019. [En línea]. Available: <https://github.com/sevamoo/SOMPY/tree/master/sompy>. [Último acceso: 15 febrero 2020].
- [3] E. Batista de Oliveira, J. E. Arce y J. C. Ferreira, «Artificial neural networks applied in forest biometrics and modeling: State of the art (January/2007 to July/2018),» *SCIELO*, vol. 25, n° 2, p. 10, 2019.
- [4] E. G. Aguilar Rivera y D. A. Dávila Garzón, «Análisis, diseño e implementación de la aplicación web para el manejo del distributivo de la Facultad de Ingeniería,» 2013. [En línea]. Available: <http://dspace.ucuenca.edu.ec/bitstream/123456789/4303/1/tesis.pdf>. [Último acceso: 08 Agosto 2018].
- [5] S. Barba Romero, «Evaluación multicriterio de proyectos, en Martínez,» *Ciencia, Tecnología y Desarrollo: Interrelaciones Teóricas y Metodológicas,*, pp. 455-507, 1994.
- [6] S. Havre, B. Hetzler y L. Nowell, «ThemeRiver: Visualizing Theme Changes over Time,» [En línea]. Available: <http://www.ifs.tuwien.ac.at/~silvia/wien/vu-infovis/references/havre-ieeeinfovis00.pdf>. [Último acceso: 25 Noviembre 2019].
- [7] N. González Benítez, V. Estrada Sentí y A. Febles Estrada, «Estudio y selección de las técnicas de Inteligencia Artificial para el diagnóstico de enfermedades,» *SCIELO*, vol. 22, n° 3, p. 6, 2018.
- [8] S. Iberdrola, «¿Somos conscientes de los retos y principales aplicaciones de la Inteligencia Artificial?,» *IBERDROLA*, 2020. [En línea]. Available:

<https://www.iberdrola.com/innovacion/que-es-inteligencia-artificial>.

[Último acceso: 05 Enero 2020].

- [9] R. Salas, «Redes neuronales artificiales,» Universidad de Valparaíso. Departamento de Computación, 1., Valparaíso, 2004.
- [10] J. Arranz de la Peña y A. Parra Truyol, «Arranz De La Peña, Jorge; Parra Truyol, Antonio. Algoritmos genéticos,» Universidad Carlos III, 2007.
- [11] E. Curiel Marín, I. Passoni, O. M. Eva M. y A. Fernández Cano, «Los mapas auto-organizados para la evaluación de la investigación de tesis doctorales: el caso de la Didáctica de las Ciencias Sociales de España,» *RELIEVE*, vol. 1, n° 2, p. 20, 2018.
- [12] G. Shelley, «RIUMA,» 01 Junio 2016. [En línea]. Available: https://riuma.uma.es/xmlui/bitstream/handle/10630/11541/S.%20Gill_TFG.pdf?sequence=1&isAllowed=y.. [Último acceso: 03 Enero 2019].
- [13] R. Martínez Pérez y E. Rodríguez Esponda, «Manual de la Metodología de la Investigación Científica,» *INFOMED*, vol. 1, n° 1, p. 74, 2016.
- [14] J. Moreno Cevallos y B. Dueñas Holguín, «Sistemas de información empresarial: la información como recurso estratégico,» *Revista Científica Dominio de las Ciencias*, vol. 4, n° 1, pp. 141 - 154, 2018.
- [15] R. Andreu, J. Ricart y J. Valor, *Estrategia y sistemas de información*. Madrid: Mc Graw Hill, 1991.
- [16] S. Havre, B. Hetzler y L. Nowell, «ThemeRiver: In Search of Trends, Patterns, and Relationships,» *Inu*, vol. 1, p. 4, 2015.
- [17] E. G. Aguilar Riera y D. A. Davila Garzon, «Análisis, diseño e implementación de la aplicación web para el manejo del distributivo de la Facultad de Ingeniería,» 2013. [En línea]. Available: <http://dspace.ucuenca.edu.ec/bitstream/123456789/4303/1/tesis.pdf>. [Último acceso: 08 Agosto 2018].

- [18] EUATM, «Introducción a la Web,» [En línea]. Available: <http://www.edificacion.upm.es/informatica/documentos/www.pdf>. [Último acceso: 2018 Agosto 2018].
- [19] V. R. Palasí Lallana, «Modelos de Desarrollo Iterativos,» *REDICCES*, vol. 4, n° 12, 2004.
- [20] S. Paloma, A. Bernal y T. Rodríguez, «Desarrollo Iterativo e Incremental,» [En línea]. Available: <https://danelly1236.files.wordpress.com/2013/12/modelo-iterativo-incrementalpresentacion2.pdf>. [Último acceso: 09 Agosto 2018].
- [21] L. A. Calvo Valverde, «Metodología iterativa de desarrollo de software para microempresas,» *Scielo*, vol. 28, n° 3, pp. 99 - 115, 2015.
- [22] S. Perez, E. Picado y G. Salgado, «Realización de un portal Web con información de los egresados de la Unan.León, Implementando software libre,» 2007. [En línea]. Available: <http://Riul.Unanleon.Edu.Ni:8080/Jspui/Bitstream/123456789/1139/1/201725.Pdf>. [Último acceso: 09 Agosto 2018].
- [23] I. 12207, «ISO/IEC 12207,» 2008.
- [24] E. M. Chango Holguin y A. G. Rivera Chasiquiza, «Sistema de gestión para la Unidad de Titulación de la Facultad de Ciencias de la Ingeniería y Aplicadas,» 2017. [En línea]. Available: <http://repositorio.utc.edu.ec/bitstream/27000/4415/1/PI-000639.pdf>. [Último acceso: 15 Agosto 2018].
- [25] E. «ECURED,» ECURED, 2020. [En línea]. Available: [https://www.ecured.cu/Lenguaje_de_programaci%C3%B3n_\(inform%C3%A1tica\)](https://www.ecured.cu/Lenguaje_de_programaci%C3%B3n_(inform%C3%A1tica)). [Último acceso: agosto 2020].
- [26] F. Pedregosa, G. Varoquaux y A. Gramfort, «Scikit-learn: Machine Learning in Python,» *JMLR*, vol. 3, n° 11, pp. 3 - 11, 2011.

- [27] C. Severance, «Python para informáticos,» 2009. [En línea]. Available: <http://do1.drchuck.net/py4inf/ES-es/book.pdf>. [Último acceso: 09 Agosto 2018]. . [Último acceso: 09 Agosto 2018].
- [28] J. Molina Ríos, N. Loja Mora, P. Zea Ordoñez y E. Loaiza Sojos, «Evaluación de los Frameworks en el desarrollo de Aplicaciones Web con Python,» *UNLA*, vol. 4, n° 4, 2016.
- [29] A. Holovaty y J. Kaplan-Moss, Holovaty, Adrian; Kaplan-Moss, Jacob. The definitive guide to Django: Web development done right, Apress, 2009.
- [30] G. Sotolongo Aguilar, M. V. Guzmán Sánchez y H. Carrillo, «VIBLIOSOM: Visualización de Información Bibliométrica mediante el Mapeo Autoorganizado,» *Redalyc*, 2011.
- [31] J. J. Montaña Moreno, «Redes Neuronales Artificiales aplicadas al Análisis de Datos,» *Scielo*, vol. 1, n° 1, p. 315, 2002.
- [32] M. Herrera Villafranca, C. Guerra Bustillo, V. Torres Cárdenas y Y. Caraballo Pérez, «Escalamiento Multidimensional y Mapas Autoorganizados para visualizar el uso de los Métodos Estadísticos no paramétricos en la rama de las Ciencias Agraria Y Biológica,» *Google Académico*, vol. 43, n° 1, pp. 51 - 56, 2012.
- [33] C. P. López y D. S. González, «Minería de datos: técnicas y herramientas,» Editorial Paraninfo, Madrid, 2007.
- [34] E. Botta-Ferret y J. Cabrera-Gato, «Minería de textos: una herramienta útil para mejorar la gestión del bibliotecario en el entorno digital,» *Acimed*, 2007.
- [35] I. Guzmán Raja, «Guzmán Raja, Isidoro, et al. Predicción de resultados empresariales versus medidas no paramétricas de eficiencia técnica: evidencia para pymes de la Región de Murcia,» Murcia, 2005.

- [36] J. Navarro, «Definición ABC,» Diciembre 2008. [En línea]. Available: <https://www.definicionabc.com/ciencia/metodologia.php>. [Último acceso: 03 Septiembre 2020].
- [37] V. Moosavi, S. Packmann y I. Valls, «GITHUB,» 2014. [En línea]. Available: <https://github.com/sevamoo/SOMPY>. [Último acceso: 15 febrero 2020].
- [38] V. Moosavi, «A Self Organizing Map (SOM) Package in Python: (SOMPY),» 18 febrero 2014. [En línea]. Available: <https://vahidmoosavi1984.wordpress.com/2014/02/18/a-self-organizing-map-som-package-in-python-sompy/>. [Último acceso: 15 febrero 2020].
- [39] M. Alaminos, A. Campos Sánchez, M. D. Caracuel, A. Rodríguez Morata, I. A. Rodríguez y }, «Modelos didácticos para el autoaprendizaje,» Actual MEd, 2009.
- [40] J. Amaya, «Toma de decisiones gerenciales: métodos cuantitativos para la administración,» Ecoe ediciones, 2010.
- [41] S. Developer, «Software Developer,» TheMushrr00m, 19 marzo 2018. [En línea]. Available: <http://laesporadelhongo.com/manejar-entornos-virtuales-en-python/>. [Último acceso: 10 marzo 2020].
- [42] A. Fernández, Python 3 al descubierto, Madrid: Alfaomega Grupo Editorial, 2013.
- [43] M. Márqués, «Base de datos,» 2009. [En línea]. Available: http://www3.uji.es/~mmarques/apuntes_bbdd/apuntes.pdf. [Último acceso: 14 04 2020].
- [44] B. Momjian, «PostgreSQL: introduction and concepts,» Addison-Wesley, vol. 162, 2001.
- [45] J. Margraf, «Self-organizing maps for sound corpus organitation,» 01 Abril 2019. [En línea]. Available: <https://www2.ak.tu->

berlin.de/~akgroup/ak_pub/abschlussarbeiten/2019/Margraf_MasA.pdf.
[Último acceso: 05 Octubre 2019].

- [46] B. Ender y E. B. Seda, «Modeling and simulating land use/cover change using artificial neural network from remotely sensing data,» *SCIELO*, vol. 25, n° 2, p. 9, 2019.
- [47] J. R. Ferreira Junior, M. Koenigkam Santos, D. Tadao Wada, A. P. Magalhães Tenório, M. H. Nogueira Barbosa y P. M. De Azevedo Marques, «Artificial intelligence, machine learning, computer-aided diagnosis, and radiomics: advances in imaging towards to precision medicine,» *SCIELO*, vol. 52, n° 6, p. 10, 2019.
- [48] A. R. Gonzales, «Gonzalez, A. R. (2012). Clasificación de datos basado en compresión,» *Revista ECIPeru*, vol. 9, n° 1, p. 6, 2012.
- [49] C. Jiménez y J. C. Anzieta Reyes, «Clasificación no supervisada de eventos sísmicos en el volcán Tungurahua - Ecuador,» Diciembre 2019. [En línea]. Available: <http://repositorio.usfq.edu.ec/handle/23000/6036>. [Último acceso: 05 Octubre 2019].
- [50] D. Ramos, R. Noriega, J. R. Laínez y A. Durango, *Curso de Ingeniería de Software: 2ª Edición.*, IT Campus Academy, 2017.
- [51] J. Castro, «Corponet sap recognized expertise,» SAP, 16 mayo 2017. [En línea]. Available: <https://blog.corponet.com.mx/estrategias-para-triunfar-con-tu-implementacion-erp>. [Último acceso: 15 enero 2020].
- [52] Anónimo, «Mundo kramer's blog,» WordPress.com, 23 mayo 2011. [En línea]. Available: <https://mundokramer.wordpress.com/2011/05/20/modelo-de-analisis-software/>. [Último acceso: 15 enero 2020].
- [53] J. Arencibia Cobas, Y. D. C. Toll Palma, J. A. Soto Pérez, D. Tamayo Ruedas, Y. Moyares Norchales y Y. Ril Gil, «no solo usabilidad: revista sobre personas, diseño y tecnología,» Universidad Internacional de la Rioja, 9 febrero 2012. [En línea]. Available:

http://www.nosolousabilidad.com/articulos/guia_ai.htm. [Último acceso: 15 enero 2020].

- [54] IONOS, «Digital Guide IONOS,» IONOS, 21 enero 2020. [En línea]. Available: <https://www.ionos.es/digitalguide/servidores/known-how/linux-vs-windows-el-gran-cuadro-comparativo/>. [Último acceso: 15 agosto 2020].
- [55] J. M. Seixas, J. Souza, A. Viera y A. Kritski, «Artificial neural network models to support the diagnosis of pleural tuberculosis in adult patients. Int J Tuberc Lung Dis,» 2013. [En línea]. Available: <http://www.ingentaconnect.com/content/iuatld/ijtd/2013/00000017/00000005/art00021?token=00581a1cffd6a264d37e41225f40384d576b4628486b253e2c49576b3427656c3c6a333f2566e4ed81d7b599>. [Último acceso: 28 Octubre 2019].
- [56] L. A. Calvo Valverde, «Metodología iterativa de desarrollo de software para microempresas».
- [57] E. Ortiz, «Los Niveles Teóricos y Metodológicos en la Institución Educativa,» *Cinta Moebio*, vol. 1, n° 43, p. 10, 2012.



ANEXOS

ANEXO I. Cuestionario de entrevista al coordinador de REDIC

UNIVERSIDAD TÉCNICA DE COTOPAXI

MAESTRÍA EN SISTEMAS DE INFORMACIÓN

Objetivo: Recolectar información relacionada a visualización de datos de producción científica en la Universidad Técnica de Cotopaxi para comprobar las necesidades de la propuesta.

Cuestionario Entrevista:

1. ¿De qué trata el sistema ECUCIENCIA?

Actualmente se ha venido desarrollando el sistema como proyectos de pregrado con alumnos de la carrera de Ingeniería en Informática, el sistema ECUCIENCIA trata sobre el manejo de datos cuantitativos de los docentes investigadores de la Universidad Técnica de Cotopaxi, es decir de sus artículos, libros o ponencias que van desarrollando y agregando a revistas científicas.

2. ¿Qué información gestiona el sistema ECUCIENCIA?

La información que gestiona es como ya se mencionó los artículos, libros y ponencias que son vinculadas a revistas científicas, lo que ayuda a los docentes de la Universidad Técnica de Cotopaxi en la toma de decisiones, la cual es mostrada en el sistema para su análisis.

3. ¿Quiénes son los usuarios que interactúan directa o indirectamente con el sistema ECUCIENCIA?

Los usuarios que interactúan directamente con el sistema son los docentes investigadores que registran sus artículos, libros y ponencias en el sistema. Además, al ser un sistema web existen usuarios externos que tendrán acceso a la visualización de los resultados.

4. ¿Cuál es la información necesaria para el mapeo auto-organizado?

Para el mapeo auto-organizado es necesario analizar los datos de artículos, libros y ponencias que se han registrado en el sistema, aquí se debe tomar en cuenta las líneas y sublíneas de investigación que rigen en la Universidad

Técnica de Cotopaxi; y los campos amplios, específicos y detallados de acuerdo a la UNESCO y CACES.

5. ¿La información se encuentra clasificada correctamente para su manejo?

La información es registrada por los docentes investigadores directamente en el sistema por lo que estaría clasificada de acuerdo a sus criterios. En este sentido está clasificada por la Facultad y Carrera a la que el docente pertenece.

6. ¿En qué lenguaje de programación fue desarrollado el sistema ECUCIENCIA y utiliza algún framework?

El lenguaje de programación con el que se ha venido desarrollando el sistema es Python, este se lo ha utilizado por su manejo de datos y librerías existentes para el desarrollo de proyectos de este tipo. Al ser un sistema web se optó por la utilización del framework Django.

7. ¿Qué gestor de base de datos es con el que se ha venido trabajando?

El sistema se ha desarrollado con el gestor de base de datos PostgreSQL, debido a su robustez en el manejo de datos y excelente acoplamiento con el framework y lenguaje de programación utilizado.

ANEXO II. Historias de Usuario

Tabla 28. Historias de Usuario

Historia de Usuario			
Número	1	Usuario:	Usuarios
Nombre de la Historia de Usuario	Gestión de la información		
Prioridad	Alta		
Responsable	Alex Geovanny Rivera Chasiquiza		
Descripción	El usuario gestionara la acción a realizar mediante filtros, los cuales ejecutarán el algoritmo para el entrenamiento.		

Elaborado por: Alex Geovanny Rivera Chasiquiza

La historia de usuario N° 1 reflejada en la Tabla 28 corresponde a la gestión de la información desde la preclasificación de la data hasta tener la data resultante.

Tabla 29. Historias de Usuario

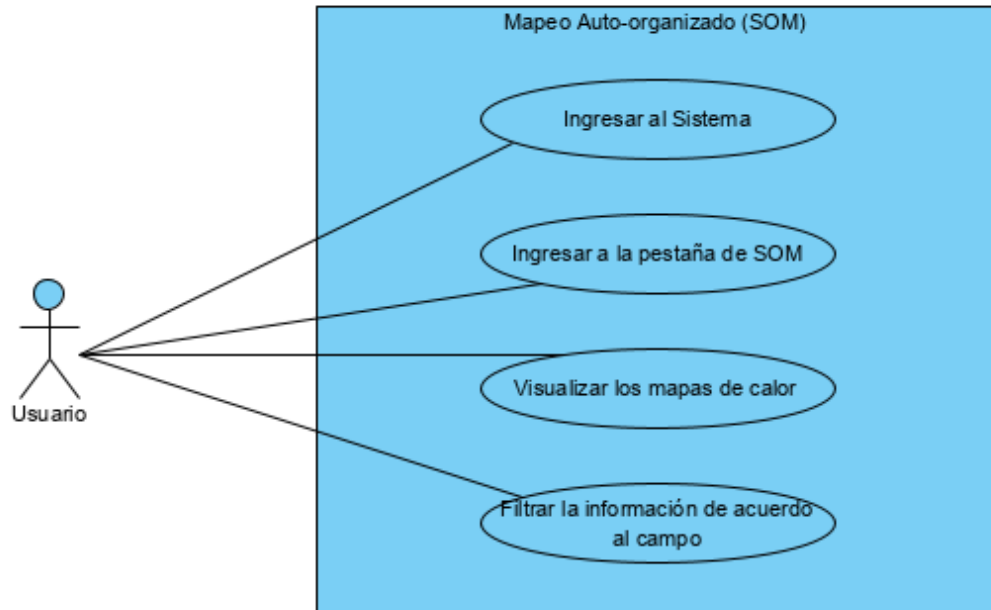
Historia de Usuario			
Número	2	Usuario:	Usuarios
Nombre de la Historia de Usuario	Visualización de la información		
Prioridad	Alta		
Responsable	Alex Geovanny Rivera Chasiquiza		
Descripción	El usuario visualizara la información solicitada, ya procesada y entrenada mediante mapas de calor.		

Elaborado por: Alex Geovanny Rivera Chasiquiza

La historia de usuario N° 2 corresponde a la visualización de la información mediante los mapas de calor, esto a partir de los datos entrenados.

ANEXO III. Diagrama de casos de uso general

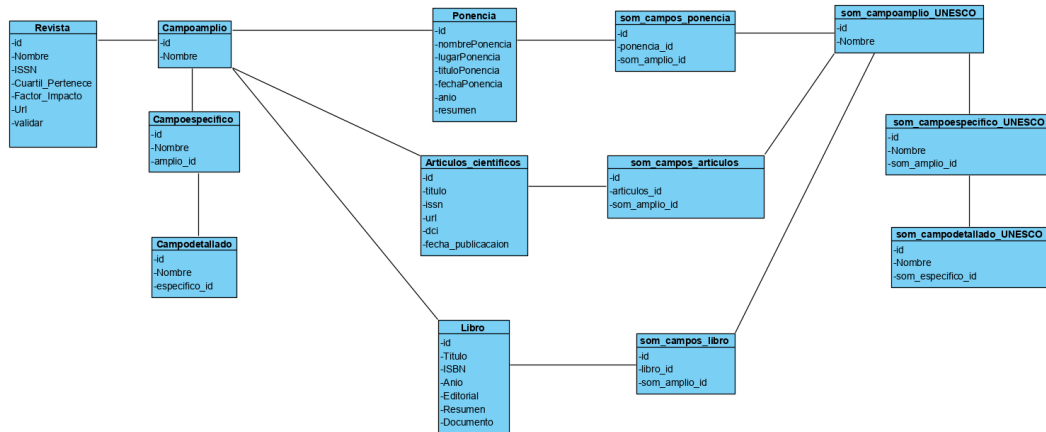
Figura 34. Diagrama de Casos de Uso General



Elaborado por: Alex Geovanny Rivera Chasiquiza

ANEXO IV. Diagrama de clases y Diagrama de secuencia

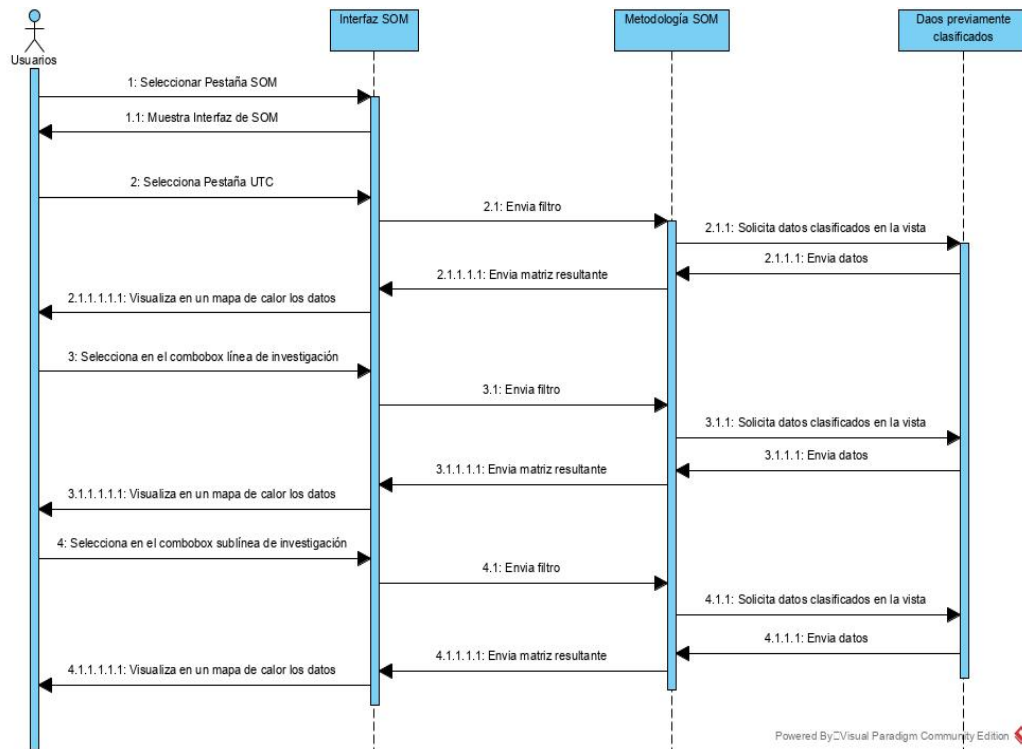
Figura 35. Diagrama de clases



Elaborado por: Alex Geovanny Rivera Chasiquiza

En la Figura 37 se puede observar el diagrama de clases que interactúa con la data necesaria para la carga de datos que serán entrenados y posteriormente visualizados.

Figura 36. Diagrama de Secuencia



Elaborado por: Alex Geovanny Rivera Chasiqiza

En la Figura 38 se puede observar el diagrama de secuencia correspondiente al caso de uso C002. Este Figura representa el proceso en general de los demás casos de uso, por lo que será el único realizado.

ANEXO V. Resultados de la implementación

Figura 37. Código View.py

```

40 IsConexion = "host=%s port=%s user=%s password=%s dbname=%s" % (IsHost, IsPort, IsUser, IsPassword, IsDBName)
41 IsOpendb = psycopg2.connect(IsConexion)
42
43 cursorca = IsOpendb.cursor()
44 cursorce = IsOpendb.cursor()
45 cursorcd = IsOpendb.cursor()
46
47 # Armaos el query de consulta a la vista del campo amplio, especifico y detallado
48
49 IsQryca = """
50     SELECT
51     count("v_CamposCaces".numcamplo) as numeroca
52     FROM
53     "v_CamposCaces"
54     GROUP BY
55     "v_CamposCaces".numcamplo;
56     """
57
58 IsQryce = """
59     SELECT
60     count("v_CamposCaces".numespecifico) as numeroce
61     FROM
62     "v_CamposCaces"
63     GROUP BY
64     "v_CamposCaces".numespecifico;
65     """
66
67 IsQrycd = """
68     SELECT
69     count("v_CamposCaces".numdetalle) as numerocd
70     FROM
71     "v_CamposCaces"
72     GROUP BY
73     "v_CamposCaces".numdetalle;
74     """

```

Elaborado por: Alex Geovanny Rivera Chasiqiza

La Figura 39 muestra un fragmento de código programado en el archivo view.py, el cual es utilizado para la extracción de la data de su respectiva vista de la base de datos PostgreSQL.

Figura 38. Código som_listar.html

```
[% block scripts %]
<!-- datatables -->
<script type="text/javascript" src="{% static 'assets/plugins/datatables/jquery.dataTables.min.js' %}"></script>
<script type="text/javascript" src="{% static 'assets/plugins/datatables/dataTables.bootstrap4.min.js' %}"></script>
<script type="text/javascript" src="{% static 'assets/plugins/datatables/responsive.min.js' %}"></script>
<script type="text/javascript" src="//ajax.googleapis.com/ajax/libs/jquery/1.8.3/jquery.min.js"></script>
<!-- load plotly.js into the page -->
<script src="https://cdn.plot.ly/plotly-latest.min.js"></script>
<!-- <script type="text/javascript" src="{% static 'assets/plugins/datatables/js/responsive.bootstrap4.min.js' %}"></script>

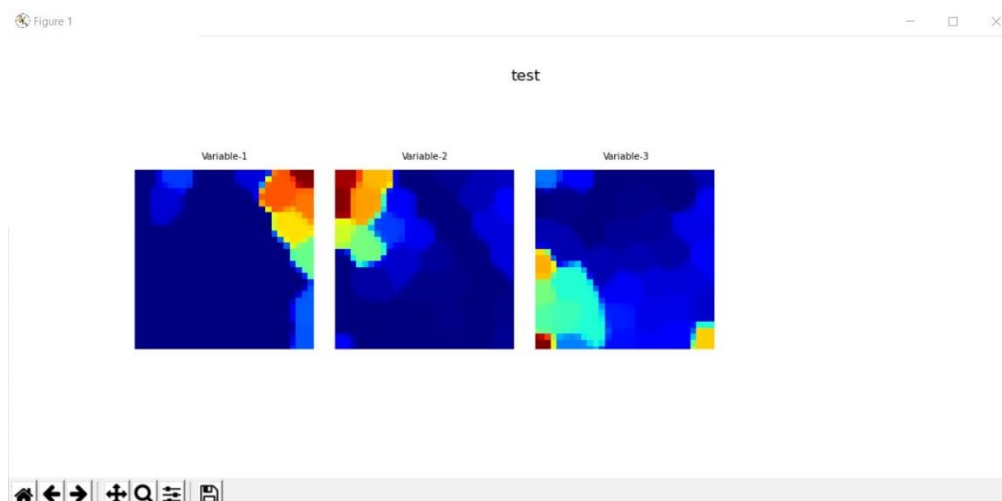
<script>
    var data = [
        {
            z: [[0.448042791455317, 2.5348648243924883, -0.07714983810280612], [-0.811712155876652758, -0.36267767571644],
            x: ['Campo Amplio', 'Campo específico', 'Campo Detallado'],
            y: ['1', '2', '3', '4', '5', '6', '7', '8', '9', '10', '11', '12', '13', '14', '15', '16', '17', '18', '19', '20'],
            type: 'heatmap',
            hoverongaps: false
        }
    ];

    Plotly.newPlot('my01v', data);
</script>
[% endblock %]
```

Elaborado por: Alex Geovanny Rivera Chasiquiza

La Figura 40 muestra un fragmento de código que ejemplifica el template con la programación en JavaScript para la carga de los resultados en el mapa de calor.

Figura 39. Mapas de calor con matplotlib

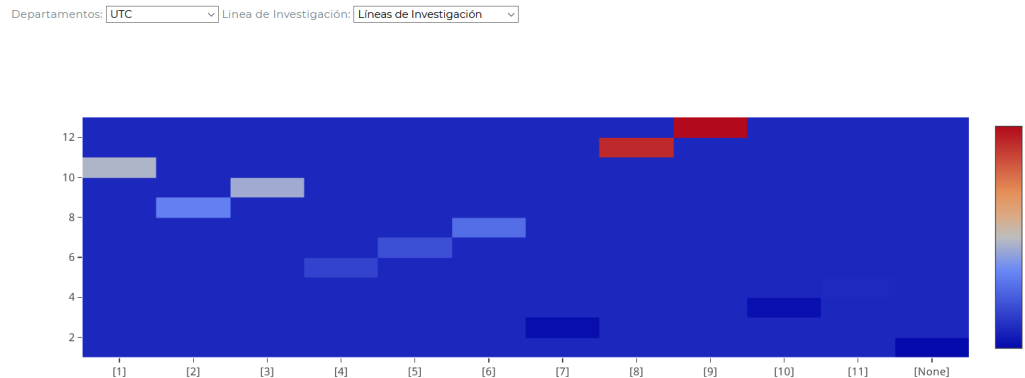


Elaborado por: Alex Geovanny Rivera Chasiquiza

En la Figura 41 se puede apreciar los resultados en mapas de calor de los campos amplios, específicos y detallados de acuerdo al CACES, la interpretación es en una

escala mínima representada con el color azul y una máxima representada por el color rojo.

Figura 40. Mapa de calor con JavaScript



Elaborado por: Alex Geovanny Rivera Chasiqiza

Figura 41. Listado de Líneas de Investigación de la Universidad Técnica de Cotopaxi

Id Nombre

- 1 TECNOLOGÍAS DE LA INFORMACIÓN Y COMUNICACIÓN (TICS)
- 2 ANÁLISIS, CONSERVACIÓN Y APROVECHAMIENTO DE LA BIODIVERSIDAD LOCAL
- 3 DESARROLLO Y SEGURIDAD ALIMENTARIA
- 4 SALUD ANIMAL
- 5 PROCESOS INDUSTRIALES
- 6 ENERGÍAS ALTERNATIVAS Y RENOVABLES, EFICIENCIA ENERGÉTICA Y PROTECCIÓN AMBIENTAL
- 7 GESTIÓN DE LA CALIDAD Y SEGURIDAD LABORAL
- 8 ADMINISTRACIÓN Y ECONOMÍA PARA EL DESARROLLO HUMANO Y SOCIAL
- 9 EDUCACIÓN, COMUNICACIÓN Y DISEÑO GRÁFICO PARA EL DESARROLLO HUMANO Y SOCIAL
- 10 CULTURA, PATRIMONIO Y SABERES ANCESTRALES
- 11 PLANIFICACIÓN Y GESTIÓN DEL TURISMO SOSTENIBLE

Elaborado por: Alex Geovanny Rivera Chasiqiza

La Figura 42 representa de manera más interactiva por medio de un mapa de calor distinto al anterior, los resultados de la data y su aprendizaje, de igual manera del CACES y sus campos amplios, específicos y detallados.

ANEXO VI. Ficha e Informe de Pruebas.

Tabla 30. *Ficha de Pruebas*

Caso de Prueba			
Caso de Uso	CU_002	Fecha Inicio:	04/05/2020
		Fecha Fin:	05/05/2020
Descripción	Caso de prueba que permite realizar los filtros de acuerdo a las líneas y sublíneas de investigación de la Universidad Técnica de Cotopaxi.		
Condiciones de ejecución	Mediante los filtros propuestos el usuario maneja de forma sencilla la búsqueda información entrenada y visualizada por mapas de calor		
Entradas	<ul style="list-style-type: none"> • El usuario selecciona la pestaña SOM • El usuario selecciona la pestaña UTC • El usuario selecciona en el combobox la línea o sublínea de investigación. • El sistema visualiza la matriz resultante del entrenamiento por medio de mapas de calor 		
Evaluación de prueba	“SUPERADO”		
Responsable	Equipo de desarrollo		

Elaborado por: Alex Geovanny Rivera Chasiquiza

En la Tabla 30 se puede apreciar la Ficha de pruebas realizada con el caso de uso CU_002, esta muestra los filtros y resultados de las líneas y sublíneas de investigación con respecto a la Universidad Técnica de Cotopaxi.

Tabla 31. Informe de Pruebas

N° Caso de Prueba	Fecha de Prueba	Resultados	
		SUPERADA	NO SUPERADA
CP001	30/04/2020	X	
	01/05/2020		
CP002	04/05/2020	X	
	05/05/2020		
CP003	06/05/2020	X	
	07/05/2020		
CP004	08/05/2020	X	
	11/05/2020		

Elaborado por: Alex Geovanny Rivera Chasiqiza

Equipo de Desarrollo

En la Tabla 31 se observa el informe de pruebas realizada durante la investigación, detallado de acuerdo a las fechas en cada caso de prueba.

ANEXO VII. Servidor de alojamiento del módulo SOM

En la actualidad la constante evolución de la big data, inteligencia artificial (IA) y machine learning ha provocado el desarrollo de software que se alinean a ellos, ya que el uso de la información impulsa a un crecimiento sostenible para las empresas. Los proyectos desarrollados con IA consumen una gran cantidad de recursos debido a los volúmenes de datos exorbitantes, es por este motivo que a continuación se detalla las tablas comparativas según [54] entre los sistemas operativos Linux y Windows. Estos datos permitirán utilizar el correcto sistema operativo como servidor para alojar el módulo propuesto de visualización de información de docentes investigadores en la Universidad Técnica de Cotopaxi mediante mapas auto-organizados (SOM).

Tabla 32. Características de los sistemas operativos Linux y Windows.

	Hosting web de Windows	Hosting web de Linux
Tipo de software	Privado	De código abierto
Servidor web	Microsoft IIS	Apache, Nginx
Lenguajes script	VBScript, ASP.NET	Perl, PHP, Python, Ruby
Bases de datos	Microsoft SQL Server, Microsoft Access	MySQL, MariaDB
Software de gestión	Plesk	cPanel, Plesk, Confixx
Otros	Exchange, aplicaciones .NET, SharePoint	WordPress, Joomla, etc.

Elaborado por: Digital Guide IONOS

Tabla 33. Ventajas e inconvenientes de Linux como sistema operativo para servidores web

Ventajas	Inconvenientes
Uso gratuito	Difícil de usar
Los administradores se benefician de las libertades de gestión del sistema	El inglés es el idioma estándar para las líneas de comandos y los mensajes del sistema
Soporta el trabajo cooperativo sin que los usuarios habituales puedan dañar el núcleo del programa	Otros programas de terceros solo pueden ser instalados por un administrador
Rara vez se ve amenazado por los cibercriminales	La portabilidad de las distribuciones de Linux no es prioritaria para muchos desarrolladores de software y de hardware

Errores de seguridad poco habituales que se solucionan rápidamente	En ocasiones, las actualizaciones son muy complejas
Pocos requisitos de hardware	No todas las versiones cuentan con asistencia a largo plazo
Función remota integrada para el control a distancia	Algunos programas profesionales no funcionan con Linux

Elaborado por: Digital Guide IONOS

Tabla 34. Ventajas e inconvenientes de Windows como sistema operativo para servidores web

Ventajas	Inconvenientes
Apto para principiantes, manejo intuitivo por medio de interfaces gráficas de usuario	Elevados costes de licencia que aumentan con cada usuario
Se puede acceder a los controladores para el hardware actual fácilmente	Fallos frecuentes de seguridad
Soporta un gran número de aplicaciones de terceros	Vulnerable a malware
Actualización de sistema sencilla y automatizada	Requiere la utilización de muchos recursos (sobre todo debido a las GUI obligatorias)
Solución de problemas técnicos por medio de la recuperación del sistema	Elevado potencial de errores de usuario
Asistencia a largo plazo garantizada	No es apto como sistema multiusuario
Posibilidad de emplear programas de Microsoft exclusivos y populares como SharePoint o Exchange	El funcionamiento del sistema propietario no es totalmente público

Elaborado por: Digital Guide IONOS

Tabla 35. Linux vs. Windows: tabla comparativa

	Windows	Linux
Costes	Costes de licencia por usuario	Sin costes de licencia; los costes de asistencia dependen de las distribuciones
Uso estándar	Interfaz gráfica de usuario	Líneas de comandos
Acceso remoto	Servidor de terminales; el cliente tiene que instalarse y configurarse	Solución integrada (terminal y shell)
Software y características	Soporta programas habituales; posibilidad de utilizar aplicaciones de Microsoft	No ofrece portabilidad para todos los programas; gran cantidad de aplicaciones disponibles
Soporte de hardware	El nuevo hardware está diseñado normalmente para los sistemas Windows	Por lo general, pueden utilizarse los controladores de hardware para las

		distribuciones de Linux más tarde
Seguridad	Elevado potencial de errores de usuario; interfaz integrada como posible punto de ataque	Los usuarios habituales no tienen acceso a los ajustes básicos del sistema; las vulnerabilidades conocidas se solucionan rápidamente
Asistencia	Asistencia a largo plazo para todas las versiones	La asistencia varía en función de la distribución y de la versión
Documentación	El sistema y sus aplicaciones están muy bien documentadas, algo que difiere de los componentes de la API y de los formatos de los datos	Se conoce el código fuente completo del sistema, las API, las bibliotecas y las aplicaciones; la mayoría de manuales y de páginas informativas están en inglés

Elaborado por: Digital Guide IONOS

Al momento de decidir por el sistema operativo adecuado para la implementación del módulo propuesto de visualización de información de docentes investigadores de la Universidad Técnica de Cotopaxi mediante mapas auto-organizados (SOM) se deben tener en cuenta varios puntos importantes, como el presupuesto disponible, nivel de conocimiento en cada sistema operativo, y en general los ya mencionados en las tablas anteriores. Lo recomendable es que independientemente del sistema operativo seleccionado para la implementación del módulo, por lo menos debe tener las siguientes características para el funcionamiento óptimo:

Tabla 36. *Requisitos mínimos del sistema operativo para implementar el módulo SOM*

Procesador	Mínimo de 1.4 GHz con arquitectura de 64 bits
Memoria RAM	Mínimo de 64 GB
Espacio en Disco	Recomendable 2 TB
Ethernet	Adaptador de Ethernet de 10/100/1000 base T Gigabit

Elaborado por: Alex Geovanny Rivera Chasiqiza