



**UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL**

FACULTAD DE INGENIERÍA

CARRERA DE INGENIERÍA EN SISTEMAS COMPUTACIONALES

TEMA:

Modelo predictivo de ocurrencia de eventos extremos de la variabilidad climática en las costas ecuatorianas basado en la detección de patrones de distribución de organismos zooplanctónicos utilizando minería de datos.

AUTOR:

Macarlupo Cajas Omayra Violeta

**Trabajo de titulación previo a la obtención del título de
INGENIERA EN SISTEMAS COMPUTACIONALES**

TUTOR:

Ing. Galo Enrique Cornejo Gómez, Mgs.

Guayaquil, Ecuador

11 de septiembre de 2019



UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL

FACULTAD DE INGENNERÍA

CARRERA DE INGENIERÍA EN SISTEMAS COMPUTACIONALES

CERTIFICACIÓN

Certificamos que el presente trabajo de titulación, fue realizado en su totalidad por **Macarlupo Cajas Omayra Violeta**, como requerimiento para la obtención del título de **Ingeniera en Sistemas Computacionales**.

TUTOR

f. _____
Ing. Galo Enrique Cornejo Gómez, Mgs.

DIRECTOR DE LA CARRERA

f. _____
Ing. Ana Isabel Camacho, Mgs

Guayaquil, a los 11 días del mes de septiembre del año 2019



UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL
FACULTAD DE INGENNERÍA

CARRERA DE INGENIERÍA EN SISTEMAS COMPUTACIONALES

DECLARACIÓN DE RESPONSABILIDAD

Yo, **Macarlupo Cajas Omayra Violeta**

DECLARO QUE:

El Trabajo de Titulación, **Modelo predictivo de ocurrencia de eventos extremos de la variabilidad climática en las costas ecuatorianas basado en la detección de patrones de distribución de organismos zooplanctónicos utilizando minería de datos** previo a la obtención del título de **Ingeniera en Sistemas Computacionales**, ha sido desarrollado respetando derechos intelectuales de terceros conforme las citas que constan en el documento, cuyas fuentes se incorporan en las referencias o bibliografías. Consecuentemente este trabajo es de mi total autoría.

En virtud de esta declaración, me responsabilizo del contenido, veracidad y alcance del Trabajo de Titulación referido.

Guayaquil, a los 11 días del mes de septiembre del año 2019

EL AUTOR

f. _____
Macarlupo Cajas Omayra Violeta



UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL

FACULTAD DE INGENNERÍA

CARRERA DE INGENIERÍA EN SISTEMAS COMPUTACIONALES

AUTORIZACIÓN

Yo, **Macarlupo Cajas Omayra Violeta**

Autorizo a la Universidad Católica de Santiago de Guayaquil a la **publicación** en la biblioteca de la institución del Trabajo de Titulación, **Modelo predictivo de ocurrencia de eventos extremos de la variabilidad climática en las costas ecuatorianas basado en la detección de patrones de distribución de organismos zooplanctónicos utilizando minería de datos**, cuyo contenido, ideas y criterios son de mi exclusiva responsabilidad y total autoría.

Guayaquil, a los 11 días del mes de septiembre del año 2019

EL AUTOR:

f. _____
Macarlupo Cajas Omayra Violeta



UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL

FACULTAD DE INGENIERIA

CARRERA DE INGENIERIA EN SISTEMAS COMPUTACIONALES

REPORTE URKUND

← → ↻ secure.orkund.com/old/view/53865977-126386-751871#q1bKLvayio7VUSrOTM/LTMtMTsxLTIWymqgFAA=

URKUND

Documento	OMAYRA_MACARLUPO_CAJAS.docx (D55364110)
Presentado	2019-09-07 11:59 (-05:00)
Presentado por	omayra_94@hotmail.com
Recibido	galo.cornejo.ucsg@analysis.orkund.com
Mensaje	OMAYRA_MACARLUPO_CAJAS.docx Mostrar el mensaje completo

0% de estas 42 páginas, se componen de texto presente en 0 fuentes.

TUTOR

f. _____

Ing. Galo Enrique Cornejo Gómez, Mgs.

AGRADECIMIENTO

Deseo expresar mis más sinceros agradecimientos a mi familia, amigos cercanos, profesor guía, y todas aquellas personas que de una u otra manera contribuyeron en la realización de esta investigación.

DEDICATORIA

El presente trabajo de investigación está dedicado a mi familia y amigos muy cercanos, que han sido un apoyo fundamental en el desarrollo y culminación de mi carrera universitaria.



**UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL
FACULTAD DE INGENNERÍA
CARRERA DE INGENIERÍA EN SISTEMAS COMPUTACIONALES**

TRIBUNAL DE SUSTENTACIÓN

f. _____

Ing. Camacho Coronel Ana Isabel, Mgs.
DIRECTOR DE CARRERA

f. _____

Ing. Castro Aguilar Gilberto Fernando Mgs.
COORDINADOR DEL ÁREA

f. _____

Ing. Celleri Mujica Mario Mgs.
OPONENTE

ÍNDICE

ÍNDICE.....	X
RESUMEN.....	XII
ABSTRACT.....	XIII
INTRODUCCIÓN.....	2
CAPÍTULO I.....	4
EL PROBLEMA.....	4
PLANTEAMIENTO DEL PROBLEMA.....	4
CAPÍTULO II.....	9
MARCO TEÓRICO.....	9
MARCO CONCEPTUAL.....	12
MARCO LEGAL.....	36
CAPÍTULO III.....	40
METODOLOGÍA DE LA INVESTIGACIÓN.....	40
METODOLOGÍA DE MINERÍA DE DATOS.....	41
METODOLOGÍA DE PROTIPADO RAPIDO.....	44
INSTRUMENTO DE RECOLECCIÓN DE DATOS.....	45
RESULTADO Y ANÁLISIS DE ENTREVISTAS.....	45
CAPÍTULO IV.....	50
PROPUESTA.....	50
SELECCIÓN DE LA HERRAMIENTA PARA IMPLEMENTACIÓN DE METODOLOGÍA KDD.....	50

IMPLEMENTACIÓN METODOLOGÍA KDD.....	51
CONCLUSIONES Y RECOMENDACIONES.....	72
BIBLIOGRAFÍA.....	74
ANEXOS.....	86

RESUMEN

El presente trabajo “**Modelo predictivo de ocurrencia de eventos extremos de la variabilidad climática en las costas ecuatorianas basado en la detección de patrones de distribución de organismos zooplanctónicos utilizando minería de datos**” tiene como objetivo obtener un modelo capaz de predecir la ocurrencia de eventos extremos climáticos como lo son los del ciclo ENOS (El Niño Oscilación Sur) a partir del comportamiento de los organismos zooplanctónicos, mediante el uso de técnicas de minería de datos, el gran volumen de datos usados pertenecen al programa permanente de Variabilidad Climática del Instituto Nacional de Pesca, quien cuenta con base de datos histórica referente al tema.

Las técnicas de minería de datos fueron ejecutadas a través del uso de la metodología KDD (Descubrimiento de Conocimiento en Base de Datos), mediante el uso de la herramienta KMINE, que es una plataforma tecnológica especializada en la ejecución de los algoritmos de minería de datos, tales como: J48 y PRISM, que son algoritmos que pertenecen a la minería de datos de tipo descriptiva, para la cual el tipo de algoritmos que se utilizan son los de clasificación, como los mencionados anteriormente.

El ciclo ENOS es un tema de gran interés, debido a que su ocurrencia trae consigo grandes consecuencias tales como: inundaciones, escasez, abundantes precipitaciones, causados por los fenómenos que pertenecen a este ciclo. Un modelo predictivo es una herramienta para toma de decisiones y para elaboración de planes que permitan mantener la integridad de toda la población.

El modelo predictivo planteado permite determinar la ocurrencia de eventos del Ciclo ENOS tales como: Fenómeno Niño, Fenómeno de la Niña o Condiciones normales, los cuales a través de factores relevantes tomados de la base de datos luego del análisis y proceso de minería, se genera a través de reglas obtenidas por medio de los algoritmos usados.

Palabras Clave: *Minería de datos, Ciclo ENOS, Metodología KDD, Organismos zooplanctónicos, modelo predictivo.*

ABSTRACT

The present work "Predictive model of occurrence of extreme events of climatic variability in Ecuadorian coasts based on the detection of distribution patterns of zooplanktonic organisms using data mining" aims to obtain a model capable of predicting the occurrence of extreme climatic events such as those of the ENSO (El Niño Southern Oscillation) cycle based on the behavior of zooplanktonic organisms, through the use of data mining techniques, the large volume of data used belongs to the permanent program of Climate Variability of the National Fisheries Institute, who has a historical database on the subject.

The data mining techniques were executed through the use of the KDD (Knowledge Discovery in Database) methodology, using the KMINE tool, which is a technological platform specialized in the execution of data mining algorithms , such as: J48 and PRISM, which are algorithms that belong to data mining of descriptive type, for which the type of algorithms used are those of classification, such as those mentioned above.

The ENSO cycle is a topic of great interest, because its occurrence brings with it great consequences such as: floods, scarcity, abundant rainfall, caused by the phenomena that belong to this cycle. A predictive model is a tool for decision making and for the elaboration of plans that maintain the integrity of the entire population.

The proposed predictive model allows to determine the occurrence of ENSO Cycle events such as: El Niño, La Niña or Normal Conditions, which through relevant factors taken from the database after the mining process and analysis, is generated through rules obtained through the algorithms used.

Key words: *Data mining, ENSO cycle, KDD methodology, Zooplankton organisms, predictive model.*

INTRODUCCIÓN

La variabilidad climática mide los rangos de variaciones del estado medio y otros estadísticos (desviación típica, la ocurrencia de extremos, etc.) del clima en todas las escalas espaciales y temporales, siempre que superen a las de los fenómenos meteorológicos. La variabilidad puede deberse a procesos internos naturales del sistema climático (variabilidad interna) o a variaciones de los forzamientos externos, sean por causa natural o antropogénica (variabilidad externa). (Agencia Estatal de Meteorología del Gobierno de España, 2016)

Uno de los factores que influyen en el clima es la temperatura de la superficie del mar en el Pacífico tropical central y oriental, la cual a través de los años ha oscilado entre condiciones extremas cálidas (El Niño) y frías (La Niña). (Pabón, 2002)

Los eventos “El Niño” y/o “La Niña”, debido a los cambios ya antes mencionados en la temperatura superficial del océano, afectan a los patrones de lluvias desde Indonesia hasta la costa oeste de América, recorriendo una distancia que cubre aproximadamente 20.000 Km. (Padilla, 2010)

Según información tomada de la página oficial del INOCAR (Instituto Oceanográfico de la Armada) dice que para una mayor comprensión de cómo se desarrollan los eventos El Niño/La Niña del ciclo ENOS (El Niño, Oscilación del Sur) a lo largo del Océano Pacífico Ecuatorial, se han definido cuatro regiones de acción. De Oeste a Este estas son: Regiones Niño 4, Niño 3, Niño 3.4 y Región Niño 1+2. (INOCAR)

Las costas de Ecuador incluido Galápagos están situadas en la Región Niño 1+2, esta región representa un indicador de los cambios inducidos por El Niño en los patrones de variabilidad de la costa del Pacífico de América del Sur. (INOCAR)

En dicha región existe una gran afluencia de corrientes en las cuales traen consigo organismos zooplanctónicos los cuales serán parte de este trabajo como objeto de estudio debido a que estos organismos no tienen movimiento propio y habitan en capas de agua con determinadas características físico- químicas, por lo que se puede conocer la preferencia o resistencia de algunas especies con respecto a los diversos parámetros ambientales: temperatura, salinidad, nutrientes (MIRÓ, 1971); se añade también que estos organismos permiten conocer el origen y desplazamiento de las corrientes y masas de agua; por lo cual su desplazamiento y distribución serán analizados. (Boltovskoy & Theyer, 1970)

La distribución de los organismos zooplanctónicos varía según diferentes factores tales como: temperatura y salinidad, para lo cual se determinaría que dichos organismos podrían cumplir un papel de indicadores de ocurrencia de eventos extremos en la variabilidad climática, debido a su reacción ante dichos factores.

Según lo escrito en el acta oceanográfica del pacífico del año 2014, dice que la sensibilidad a las variaciones del ambiente marino por parte de las estas especies, convierten a estos organismos en herramientas vivas de detección de cambios del ecosistema marino en un momento dado. (Luzuriaga, 2014)

En Ecuador el encargado de realizar el estudio de estos organismos en las zonas costeras del país es el Instituto Nacional de Pesca (INP), dicha institución realiza investigaciones relacionadas con procesos oceanográficos asociados a estas especies y a los eventos extremos ya antes mencionados, para la emisión de informes mensuales, por medio del procesamiento de sus datos utilizando técnicas estadísticas, para publicación de boletines informativos en conjunto con otras instituciones especializadas en el tema.

Para el desarrollo de este trabajo el INP ha proporcionado datos de sus investigaciones de los últimos diez años, con el fin de modificar el proceso estadístico habitual mediante la generación de un modelo predictivo de ocurrencia de eventos extremos de la variabilidad climática en

las costas ecuatorianas basado en la detección de patrones de distribución de organismos zooplanctónicos por medio del uso de técnicas de minería de datos.

CAPÍTULO I

EL PROBLEMA

PLANTEAMIENTO DEL PROBLEMA

Ubicación del Problema en un Contexto

Los eventos extremos de la variabilidad climática que se abordarán en este trabajo están relacionados al ciclo ENOS, el cual consta de 2 fases: la cálida, que es conocida como el Fenómeno del Niño y la fase fría o también llamada Fenómeno de La Niña; la ocurrencia de estos fenómenos trae grandes consecuencias para el Ecuador. Estos eventos anómalos no tienen periodos fijos de ocurrencia, y debido a esto, su estudio y monitoreo es constante para mantener informada a la población y sobre todo a las autoridades, para elaborar planes de contingencia ante dichos sucesos.

Los fenómenos ya antes mencionados, están asociados a la parte oceanográfica y atmosférica del ecosistema, sin embargo en este trabajo el análisis estará enfocado en los aspectos oceanográficos, los cuales involucran las masas de agua, las corrientes, los componentes y la presencia de organismos pertenecientes a la familia del zooplancton, quienes podrían determinar la ocurrencia de dichos eventos extremos, debido a características específicas de estos organismos. En la actualidad en Ecuador el análisis y estudio realizado de estos organismos no son tomados como características determinantes en la ocurrencia de estos eventos.

La zona costera del Ecuador es el lugar en donde se realizan los trabajos de campo con respecto a este tema, abarca la zona costera de

Manabí hasta la parte sur que es la provincia de El Oro. En donde investigadores encargados de estos estudios hacen el levantamiento de la información por medio de toma de muestras para su posterior análisis.

La ocurrencia de los eventos extremos de la variabilidad climática asociados al ciclo ENOS están directamente conectados a las corrientes marinas presentes en los ecosistemas marinos costeros, en los cuales habitan diferentes grupos de especies, entre los cuales están los organismos zooplanctónicos, importantes en este tema, debido a que son parte de las masas de aguas que se trasladan con las corrientes.

Cuando estos eventos suceden, las consecuencias son grandes en diferentes aspectos para el país, tanto económicas como sociales, porque traen consigo abundantes precipitaciones, con ello desbordamiento de ríos, o en su defecto en el ciclo opuesto trae sequías y ausencia de especies; ambos extremos afectan de gran manera en la zona productiva del país, lo cual está asociado directamente con la economía, y por otro lado también devastaciones para habitantes lo cual también afecta directamente a la población.

Delimitación del Problema

El modelo predictivo propuesto será generado con datos de estudios de los últimos diez años de investigaciones y trabajos del INP, los cuales están enfocados en los organismos zooplanctónicos presentes en las zonas costeras del Ecuador.

En base a los estudios del INP, se pretende detectar los patrones de distribución de los organismos zooplanctónicos mediante el uso de técnicas de minería de datos, para determinar la ocurrencia de eventos extremos de la variabilidad climática, que para el caso de este trabajo se centrarán en el ciclo ENOS que comprende las dos fases: El Niño/La Niña.

Formulación del Problema

En la actualidad en Ecuador los organismos zooplanctónicos no son tomados como bioindicadores de eventos extremos de la variabilidad

climática, solo son parte del estudio como información complementaria, pero no como factores determinantes para predicciones.

El INP según sus estudios e investigaciones realizadas indica que las condiciones oceanográficas del ecosistema marino-costero han sido muy irregulares durante estos últimos años de estudio, presentando desde condiciones normales hasta eventos extremos de la variabilidad climática, que quizás sería uno de los factores que incidieron en los patrones de distribución y abundancia de los organismos zooplanctónicos y que posiblemente serían indicadores de corrientes y procesos oceanográficos.

Por esta razón el presente estudio se basará en el uso de las técnicas de minería de datos para demostrar mediante patrones de distribución de dichos organismos que son indicadores de ocurrencia de eventos extremos tales como el fenómeno del Niño (condiciones extremas de temperaturas del agua cálidas) y la Niña (condiciones extremas de temperaturas del agua frías).

Los resultados sobre la abundancia de los organismos del zooplancton han sido documentados a través de investigaciones realizadas en el INP e INOCAR, en la actualidad estos estudios continúan, para lograr avances en el conocimiento de los cambios que se dan en el ecosistema.

Evaluación del Problema

Para la evaluación del problema se tomará en cuenta las siguientes características:

Delimitado: El problema tal y como se puede identificar en su descripción define claramente el lugar en el cual está situado y las variables que estarán involucradas en el mismo. Esta característica es importante debido que el problema al estar correctamente definido, permite ejecutar o desarrollar la solución adecuada o propuesta de mejora.

Relevante: El problema planteado es considerado relevante ya que es de interés colectivo, debido a que involucra las investigaciones realizadas por el INP sobre los fenómenos El Niño/La Niña con los cuales se realizan

informes y boletines para exponer a las autoridades pertinentes y a su vez a la comunidad para tomar medidas preventivas o de contingencia ya que sus efectos son grandes consecuencias para los mismos, lo cual lo convierte en un tema de total interés y de alto grado de relevancia.

Factible: Se considera factible por el acceso a los datos correspondientes a la base histórica proporcionada por el INP, lo que hace que esta propuesta se pueda realizar sin dificultades y con los recursos necesarios disponibles.

OBJETIVOS

OBJETIVO GENERAL

Generar modelo predictivo de ocurrencia de eventos extremos de la variabilidad climática en las costas ecuatorianas basado en la detección de patrones de distribución de organismos zooplanctónicos utilizando técnicas de minería de datos.

OBJETIVOS ESPECÍFICOS

- Realizar un diagnóstico de la distribución de organismos zooplanctónicos que identifican eventos oceánicos asociados con el ciclo ENOS (El Niño/Oscilación del Sur).
- Diseñar la vista minable para la revisión de posibles escenarios respecto al caso de estudio, que permita analizar de manera predictiva el comportamiento de resultados.
- Evaluar los resultados obtenidos a partir de la aplicación de las técnicas de minería de datos, que facilite el diseño de un modelo predictivo.
- Generar el modelo predictivo con empleo de algoritmos necesarios para detectar patrones.

ALCANCES DEL PROBLEMA

El área de estudio de esta investigación se ubica frente de la costa ecuatoriana, desde la parte central de la provincia de Manabí hasta la parte sur de El Oro a 10 millas costa afuera. El levantamiento de la información de la base de datos que será utilizada, se realizó durante un ciclo hidrológico

que representaba tanta época seca (septiembre a diciembre) y época de lluvia (marzo –abril) durante el periodo 2008-2018.

Se realizará un análisis de la información provista por el INP utilizando la metodología de minería de datos KDD (Knowledge Discovery in Databases) para encontrar patrones de distribución de organismos zooplanctónicos con diferentes algoritmos de validación para luego con un mayor volumen de información para poder comprobar que dichos organismos influyen en la ocurrencia de eventos extremos en la variabilidad climática en las costas ecuatorianas y con esto generar el modelo predictivo.

JUSTIFICACION E IMPORTANCIA

La minería de datos en la actualidad está inmersa en diferentes ramas de la ciencia, debido a su aporte en los temas de patrones, predicciones, que son de gran importancia para toma de decisiones y planes de prevención en caso que fuera necesario.

En el campo medio ambiental la minería de datos brinda las posibilidades de crear modelos que permiten comprender el comportamiento de los ecosistemas y con lo cual se pueden encontrar más posibilidades y forma de conservación de los mismos.

Como la minería de datos está enfocado a la detección de patrones, mediante el proceso de grandes cantidades de datos, y contribuye generando modelos predictivos, los cuales aportan en el tema de fenómenos naturales para los cuales con la ayuda de los modelos generados por medio de las técnicas de minería de datos y la detección de patrones, se puede realizar planes de prevención y contingencia para este tipo de eventos.

Por lo cual se espera que el modelo predictivo planteado cumpla esta función y sea parte de la toma de decisiones para la ocurrencia del ciclo ENOS.

HIPÓTESIS O PREGUNTA DE INVESTIGACIÓN

Un modelo predictivo de ocurrencia de eventos extremos de la variabilidad climática mediante la detección de patrones de distribución de

los organismos zooplanctónicos permite determinar o predecir la ocurrencia de dichos eventos.

VARIABLES DE LA INVESTIGACIÓN

- **Variable independiente:** modelo predictivo
- **Variable dependiente:** patrones de distribución de los organismos zooplanctónicos en las costas del Ecuador.

CAPÍTULO II

MARCO TEÓRICO

El mar que se encuentra situado en las costas del Ecuador tiene afluencia y concurrencia de diferentes masas de agua, y con ellas las corrientes que provienen de diferentes lugares, debido a que frente a las costas ecuatorianas se encuentran estas mezclas de corrientes y masas de aguas, las especies tales como los indicadores biológicos en esta zona son frecuentes; lo que hace importante su estudio ya que con ello se puede distinguir las diversas masas de aguas que están presentes y las características particulares de cada una. El estudio del comportamiento de las especies zooplanctónicas es importante para monitorear los procesos y ocurrencia de eventos oceanográficos tales como los que son parte de este trabajo: el fenómeno de “El Niño” y “La niña”, ya que un constante estudio ayudaría a posibles predicciones y con las cuales se podría alertar a los habitantes y a los encargados de este tipo de desastres y fenómenos para que se pueda realizar una contención o prevención oportuna, tratando siempre de evitar consecuencias lamentables para la comunidad.

La dirección, desplazamiento y características físico-químicas de las corrientes oceánicas, inciden en los cambios de la biodiversidad marina frente a Ecuador; la sensibilidad a las variaciones del ambiente de especies planctónicas, las convierten en herramientas vivas de detección de cambios del ecosistema marino en un momento dado. (Luzuriaga-Villareal, 2015)

En cada tipo de agua hay especies marinas que toleran un amplio rango de temperatura (euritermos) y otras toleran un pequeño rango de temperatura (estenotermos), así como otras especies toleran también un amplio rango de salinidad (Eurihalinos) y otras soportan un pequeño rango de salinidad (Estenohalinos), conocer el tipo de agua o masa de agua en que prefieren vivir y sus rangos de tolerancia a la temperatura y salinidad, es básico para interpretar con la presencia de cada especie, el tipo de masa de agua que está presente y conocer mejor el ambiente marino en base a estas especies bioindicadoras. (Cruz, 2012)

Todos los estudios realizados referente al tema de organismos zooplanctónicos como bioindicadores de ocurrencia de eventos extremos de la variabilidad climática, se los ha realizado en base a un trabajo de campo y toma de muestras de aguas para los análisis especializados de cada uno; para lo cual los datos obtenidos de dichos análisis se almacenan en bases de datos históricas que son usadas para publicaciones, informes y boletines informativos por parte de las Instituciones encargadas (el INP y el INOCAR).

La forma habitual del procesamiento de los datos obtenidos para generación de publicaciones es por medio de técnicas estadísticas convencionales por las cuales se muestran resultados de distribución, abundancia, promedios, varianzas, entre otros datos dependiendo el tipo de estudio e informe que se este generando.

Este tipo de estudio genera gran volumen de datos, los cuales mediante técnicas de minería de datos podrían ser procesados de manera eficaz, y por medio de lo cual se podría obtener los patrones que serán parte del análisis y con esto se puede dar una pauta para la toma de decisiones a nivel nacional con respecto a la ocurrencia de los fenómenos antes mencionados.

La minería de datos utiliza el análisis matemático para deducir los patrones y tendencias que existen en los datos. Normalmente, estos patrones no se pueden detectar mediante la exploración tradicional de los datos porque las relaciones son demasiado complejas o porque hay demasiados datos.

Estos patrones y tendencias se pueden recopilar y definir como un modelo de minería de datos. (Microsoft, 2019). Por tanto, los datos son el medio o la base para llegar a conclusiones y transformar estos datos en información relevante. (Ribas, 2018)

La información tomada de la presentación “Extracción de Conocimiento en Bases de datos de Ingeniería del Software – T.2 Minería de Datos y Extracción de Conocimiento de Bases de Datos” escrita por José Hernández y César Ferri (2018) ,resalta las siguientes diferencias entre la minería de datos y los métodos estadísticos:

- *Diferencia clara con métodos estadísticos:* la estadística se utiliza para validar o parametrizar un modelo sugerido y preexistente, no para generarlo.
- *Diferencia sutil “Análisis Inteligente de Datos”* (IDA, del inglés Intelligent Data Analysis) que correspondía con el uso de técnicas de inteligencia artificial en el análisis de los datos.

Para el uso de las técnicas de minería de datos existe una metodología llamada: “Descubrimiento de Conocimiento a partir de Bases de Datos” (KDD, del inglés Knowledge Discovery from Databases). “proceso no trivial de identificar patrones válidos, novedosos, potencialmente útiles y en última instancia comprensibles a partir de los datos”. Fayyad et al. 1996 (Orallo & Ramírez, 2018)

En el presente trabajo para el cumplimiento de su objetivo la detección de patrones de distribución de los organismos zooplanctónicos es fundamental, para lo cual según la publicación de José Hernández y César Ferri (2018) sobre Minería de Datos y Extracción de Conocimiento de Bases de Datos, siguiendo el proceso de la metodología KDD, ya antes mencionada, dice que:

Una vez recogidos los datos de interés en un almacén de datos, un explorador puede decidir qué tipos de patrón quiere descubrir; el tipo de conocimiento que se desea extraer va a marcar claramente la técnica de

minería de datos a utilizar. Según como sea la búsqueda del conocimiento se puede distinguir:

Directed data mining: se sabe claramente lo que se busca; generalmente predecir unos ciertos datos o clases.

Undirected data mining: no se sabe lo que se busca, se trabaja con los datos.

En el primer caso, los propios sistemas de minería de datos se encargan generalmente de elegir el algoritmo más idóneo entre los disponibles para un determinado tipo de patrón a buscar.

Para mayor comprensión del presente **MARCO CONCEPTUAL** trabajo a continuación se detallará los conceptos e importancia de palabras poco convencionales para los lectores de este documento; destacando su rol en el cumplimiento del objetivo del mismo; y dando a conocer características importantes de cada término; que aportan en esta investigación.

Los organismos zooplanctónicos y especies indicadoras

Organismos zooplanctónicos: La palabra Zooplancton viene del griego *Zoo* -que significa animal- y *Planktos* -que significa vagando-, este término se refiere a un grupo de animales que divagan en la columna de agua y que no tienen la capacidad de natación suficiente para contrarrestar las corrientes. El zooplancton es muy diverso y su composición y abundancia están influenciadas principalmente por sus migraciones verticales (patrón de desplazamiento diario), agregación, épocas climáticas y además por factores antropogénicos y naturales como salinidad, temperatura, entre otros. (Tobías, y otros, 2014)

Existen dos grupos de zooplancton: el holoplancton, que incluye a organismos que siempre son parte del plancton, y el meroplancton, donde los individuos son plancton sólo en una etapa de su vida. (Coto, 2015)

Debido a la abundancia de ambos tipos y su presencia en profundidades variables, el zooplancton se utiliza para evaluar la transferencia de energía en los niveles tróficos secundarios. El zooplancton se alimenta de fitoplancton –plantas microscópicas– y facilitan la conversión de material vegetal en tejido animal y, a su vez, constituyen el alimento básico para animales de niveles tróficos superiores, incluidos los peces, particularmente sus larvas. (EL COLEGIO DE LA FRONTERA SUR, 2017)

Los organismos zooplanctónicos en su mayoría son organismos microscópicos que cumplen un importante rol en los océanos, están ligados directamente a la biodiversidad de especies en los ecosistemas acuáticos debido a que forman parte de la cadena alimenticia marina, que es llamada red trófica, cumpliendo su función de ser alimento de otras especies y no solo eso, sino que su comportamiento también está muy relacionado con el medio ambiente en general.

En el desarrollo de este trabajo el estudio de estos organismos es fundamental, ya que a partir de ellos, del comportamiento en su hábitat, pueden ser indicadores claves de sucesos en el ecosistema en general.

Niveles Tróficos

Los niveles tróficos son la clasificación por grupos de especies que conforman la cadena alimenticia o también llamada red trófica, en donde se encuentran como primer nivel el grupo de los productores, en el segundo nivel se encuentra el grupo de los consumidores que pueden ser primarios, secundarios, terciarios, etc, y como último y tercer nivel se encuentran depredadores o descomponedores; cada uno de estos niveles cumple una función específica en el medio en el cual se desarrollan.

En el ecosistema marino los niveles tróficos están distribuidos de la siguiente manera: en el nivel más bajo están los productores (fitoplancton, organismos vegetales acuáticos), en el segundo nivel se encuentran los consumidores primarios (zooplancton), en el tercer nivel están los consumidores secundarios y terciarios (especies de peces) y como último

nivel están depredadores o descomponedores (ballenas, tiburones, ser humano).

En la figura 1 se muestra los niveles tróficos ya antes mencionados y las especies que los conforman a cada uno en la red trófica o también llamada cadena alimenticia.

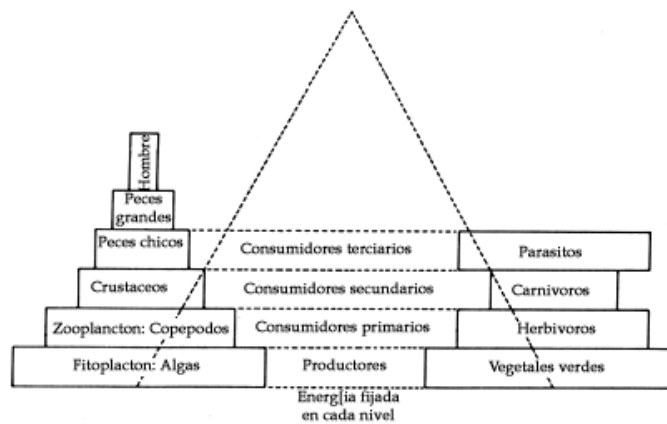


Figura 1. Niveles tróficos, tomado de (Cifuentes, Torres-García, & Mondragón, 1997)

Especies indicadoras

Se definen como aquellas que por sus características (sensibilidad a las perturbaciones ambientales, distribución, abundancia, dispersión, éxito reproductivo, entre otras) pueden ser usadas como estimadoras del estatus de otras especies o condiciones ambientales de interés que resultan difíciles, inconvenientes o costosas de medir directamente. (Heink & Kowarik, 2010)

Los bioindicadores en un estricto sentido, proporcionan información cualitativa de la salud del ambiente a través de su presencia/ausencia o mediante cambios en su abundancia. (Zuarth, Vallarino, Jiménez, & Pfeng, 2014)

Por tanto, tomando en cuenta los conceptos de bioindicadores, los organismos zooplanctónicos cumplen con algunas de las características que debe tener un bioindicador; debido al comportamiento que estos organismos tienen dentro su ecosistema.

La presencia, abundancia o escasez de los mismos son indicadores de ciertas condiciones en las corrientes del océano, por lo cual esto informa sobre el estado del ecosistema marino y debido a su escasa movilidad de manera autónoma, es un informante de que tipos de corrientes están presentes en las costas ecuatorianas en donde es el área de estudio.

Los diferentes cuerpos de agua, ya sean continentales, costeros o marinos presentan características físicas y químicas a las que se les ha denominado con el término común de calidad del agua, las cuales permiten la presencia saludable de una amplia variedad de organismos que habitan en ella; sin embargo, dadas las diversas actividades antropogénicas, dicha calidad se vuelve deletérea para los organismos que son sensibles, convirtiéndose en bioindicadores. (Zuarth, Vallarino, Jiménez, & Pfeng, 2014)

Eventos oceanográficos extremos y sus características determinantes

Para poder definir y comprender que es un evento oceanográfico primero se debe saber que es la oceanografía ya que parte de esta definición, entonces;

¿Qué es la oceanografía?

Oceanografía es la ciencia que estudia los mares y océanos, sus elementos (aguas, fondos, organismos vivos) y la dinámica del sistema. (Centro Regional de Información sobre Desastres (CRID), 2000)

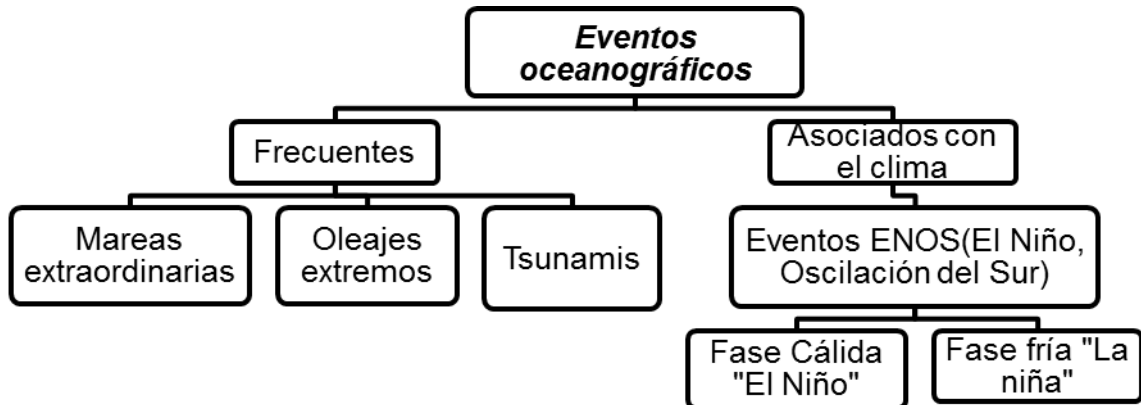
Eventos oceanográficos extremos

Tal y como dice su definición la oceanografía es una ciencia que aporta con estudios del océano y mares dichos estudios e investigaciones están enfocados principalmente a los eventos oceanográficos extremos, de los cuales se puede decir que:

Los eventos oceanográficos están asociados a las corrientes y fuerzas de las cuales es parte el agua del mar, y por lo cual está en constante movimiento, lo cual es importante para muchas de las especies del ecosistema marino, estos movimientos son los que ocasionan los oleajes, lo

mismos que crean los eventos oceanográficos más frecuentes y conocidos como: mareas extraordinarias, tsunamis, oleaje extremo. (República de El Salvador en la América Central, 2011)

Como ya fueron mencionados existen eventos oceanográficos



extremos frecuentes, pero también existen los asociados al clima, de estos últimos se detallará su explicación debido a que están involucrados y tienen un papel muy importante en este trabajo.

Gráfico 1. Clasificación de Eventos oceanográficos. Autora

Los eventos oceanográficos asociados al clima se los conoce como el ciclo ENOS, que tiene dos fases: la cálida (fenómeno de El Niño) y la fría (fenómeno de La Niña) ambas están directamente asociadas con las corrientes del océano, su temperatura y salinidad. Ambas fases serán detalladas a continuación:

Fenómeno del Niño

El Niño, corresponde a un evento climático natural que se desarrolla en el océano Pacífico ecuatorial central, la fase cálida de ENOS conocida como El Niño se manifiesta, principalmente, por un aumento de la Temperatura Superficial del Mar (TSM) y una disminución de los vientos alisios en el lado este del océano Pacífico. Estas condiciones anómalas generan fuertes precipitaciones y cambios notables en el clima y las pesquerías, tanto en los países ribereños del Pacífico sudoriental, como en otras partes del mundo. (Maturana, Bello, & Manley, 2004)

Hay una serie de índices que nos muestran la presencia de anomalías en la naturaleza que conllevan la formación de Los Niños. Los podemos dividir en índices subjetivos e índices objetivos. Los segundos se relacionan con la profundidad de la termoclina (la capa de mezcla de las aguas frías con las cálidas en el mar), la convección en la atmósfera tropical (que produce actividad nubosa y lluvias sobre las aguas del océano); la radiación de onda larga saliente (que se manifiesta en la cantidad de la cobertura nubosa; se mide en los satélites); el balance de la presión atmosférica entre el sistema de alta presión del Pacífico Sur y un sistema de baja presión cerca a Indonesia que permite el cálculo del denominado índice de Oscilación del Sur-IO; la intensidad de los vientos alisios en varias capas de la atmósfera y en superficie y las anomalías de las temperaturas superficiales del Océano Pacífico. (Henriquez-Daza, 2001).

El fenómeno del niño está marcado por el aumento de temperatura superficiales del mar, y con ello suben los niveles de las aguas, se caracteriza por abundantes precipitaciones, lo cual ocasiona grandes desastres en un país, pero en ciertas zonas que son minoría le favorece debido a la sequía permanente en dichos lugares; en la figura 2 se muestra las diferencias en el aumento de la temperatura en condiciones del fenómeno del niño y en condiciones normales

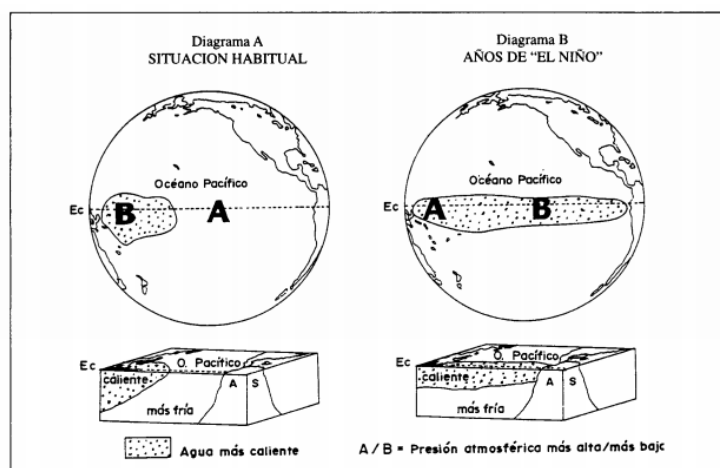


Figura 2. Condiciones de la temperatura superficial del mar en época normal y en época "El Niño". Tomado de (Espinoza-A, 1996)

Fenómeno de La Niña

El término La Niña (Fenómeno frío del Pacífico) corresponde a la aparición irregular de aguas superficiales y subsuperficiales más frías que lo normal en los sectores central y oriental del Océano Pacífico tropical. El enfriamiento de la superficie del mar cubre grandes extensiones y por su magnitud afecta el clima en diferentes regiones del planeta, entre ellas el territorio de Colombia, que se halla localizado en el borde nororiental de la cuenca. La región donde tiene desarrollo el fenómeno La Niña es la cuenca del Océano Pacífico tropical. (Pabón-Caicedo & Montealegre-Bocanera, 1998)

Aunque las condiciones La Niña son cercanamente inversas a las que se presentan durante El Niño, las anomalías negativas de temperatura superficial del mar durante la fase fría no son equivalentes o simétricamente inversas a las anomalías positivas que se registran durante los eventos cálidos. No obstante, el hecho de que de alguna manera sean condiciones contrarias a las del fenómeno El Niño, sirvió de base para que a finales de los 80s se le empezara a denominar con el nombre con el que hoy se le conoce en la comunidad científica: La Niña. (Montealegre-Bocanera, 2007)



Figura 3. Fenómeno de La niña en el Océano Pacífico. Tomado de (Viana-Pardo, 2011)

Tal y como se puede ver en la figura 3 el fenómeno de la niña tiene que ver con las corrientes de aguas frías, que también son conocidas como corrientes del Humboldt que van desde la costa occidental a la costa oriental.

Este fenómeno no tiene una característica fija de efectos en cada país, ya que en cada uno muestra sus efectos de diferentes maneras, en ciertos lugares produce sequía y en otros como en el norte produce abundantes lluvias. Este fenómeno tiene periodos irregulares ya que puede durar de

meses a años y puede repetirse el ciclo de manera constante o periódica después de muchos años.

Dado que los fenómenos El Niño y La Niña están asociados con el aumento o disminución anómalo de la temperatura superficial del mar, uno de los indicadores más utilizados para hacer seguimiento a estos fenómenos es el Índice del Niño Oceánico (ONI, por sus siglas en inglés), desarrollado por la National Oceanic and Atmospheric Administration (NOAA). Este índice es calculado a partir de mediciones de la temperatura superficial del mar en el sector central del Pacífico tropical frente a las costas de Suramérica. (Comunidad Andina, 2009)

El Índice Oceánico El Niño (ONI), que es el índice standard que usa la NOAA para identificar los fenómenos de calentamiento (El Niño) y enfriamiento (La Niña) en el Pacífico tropical, es la media de tres meses consecutivos de la anomalía de la temperatura de la superficie del océano en el sector 3.4, que es el comprendido entre los paralelos 5°N y 5°S y los meridianos 120°O y 170°O. Cuando este índice es superior a 0,5 °C durante cinco meses consecutivos estamos en El Niño, y cuando es inferior a 0,5 °C estamos en La Niña. La anomalía tiene como base el período 1971 – 2000. (Pisico, 2010)

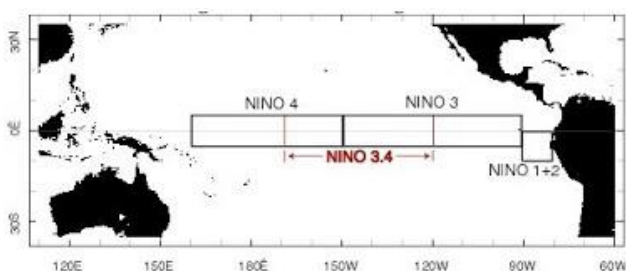


Figura 4. Regiones de acción ENOS. Tomado de (Pisico, 2010)

En la figura 4 se puede observar la división de las regiones de acción ENOS, que son: Región niño 4, Región niño 3, Región niño 3.4 y Región niño 1+2; la mayoría de predicciones sobre estos fenómenos se los realiza en la Región Niño 3.4 que es donde se encuentra EEUU; por lo cual el ONI es calculado de esa región; Ecuador se encuentra en la región niño 1+2; y los

estudios que se realizan aquí toman como referencia la región 3.4 y hace comparaciones para determinar que tan diferentes o iguales están las condiciones en cada región.

El ONI es publicado en la página oficial del NOAA en donde periódicamente es actualizado y se encuentra el registro histórico del mismo; ya que en la región 3.4 donde es calculado ocurren las corrientes originarias del ciclo ENOS, a continuación en la figura cuatro se muestra la tabla oficial

2008	-1.6	-1.4	-1.2	-0.9	-0.8	-0.5	-0.4	-0.3	-0.3	-0.4	-0.6	-0.7
2009	-0.8	-0.7	-0.5	-0.2	0.1	0.4	0.5	0.5	0.7	1.0	1.3	1.6
Año	DJF	JFM	Fma	MAM	AMJ	MJJ	JJA	Jas	ASO	HIJO	OND	NDJ
2010	1.5	1.3	0.9	0.4	-0.1	-0.6	-1.0	-1.4	-1.6	-1.7	-1.7	-1.6
2011	-1.4	-1.1	-0.8	-0.6	-0.5	-0.4	-0.5	-0.7	-0.9	-1.1	-1.1	-1.0
2012	-0.8	-0.6	-0.5	-0.4	-0.2	0.1	0.3	0.3	0.3	0.2	0.0	-0.2
2013	-0.4	-0.3	-0.2	-0.2	-0.3	-0.3	-0.4	-0.4	-0.3	-0.2	-0.2	-0.3
2014	-0.4	-0.4	-0.2	0.1	0.3	0.2	0.1	0.0	0.2	0.4	0.6	0.7
2015	0.6	0.6	0.6	0.8	1.0	1.2	1.5	1.8	2.1	2.4	2.5	2.6
2016	2.5	2.2	1.7	1.0	0.5	0.0	-0.3	-0.6	-0.7	-0.7	-0.7	-0.6
2017	-0.3	-0.1	0.1	0.3	0.4	0.4	0.2	-0.1	-0.4	-0.7	-0.9	-1.0
2018	-0.9	-0.8	-0.6	-0.4	-0.1	0.1	0.1	0.2	0.4	0.7	0.9	0.8

del ONI en los últimos 10 años.

Figura 5. Anomalías mensuales de la Temperatura Superficial del Mar (TSM).

Tomado de (NOAA, s.f.)

Según lo publicado en los registros oficiales del NOAA la tabla mostrada en la figura 5 se interpreta de la siguiente manera:

Los períodos cálido (rojo) y frío (azul) se basan en un umbral de +/- 0.5 ° C para el Índice de Niño Oceánico (ONI). (NOAA, s.f.)

Basado en el ONI , sus condiciones normales y eventos anomalos, Soren Vangni Bautista Cabezas (2015) clasifica estos eventos de la siguiente manera:

Escenario	Anomalías (grados centígrados)
Niño moderado-fuerte	mayor a (+)1
Niño débil	(+) 0.5 a (+) 1
Normal	(+/-) 0.5
Niña débil	(-) 0.5 a (-) 1
Niña moderado-fuerte	mayor a (-)1

Figura 6. Escenarios de acuerdo a las anomalías presentes en la TSM.
Tomado de (Bautista-Cabezas, 2015)

Es así como según esa clasificación existen eventos con mayor grado de ocurrencia que otros, teniendo características más notorias que las de menor grado; sabiendo esto se ha identificado la ocurrencia de estos fenómenos en el país según su grado de afectación, débil o fuerte-moderado, lo cual se muestra en la figura 7 a continuación.

Diciembre/Enero/Febrero	Años
Niño moderado o fuerte	1983, 1987, 1992, 1998, 2003, 2010
Niño débil	1998, 1995, 2005, 2007
Años neutros	1981, 1982, 1986, 1990, 1991, 1993, 1994, 1997, 2002, 2004
Niña Débil	1984, 1985, 1996, 2001, 2006, 2009
Niña moderada o fuerte	19989, 1999, 2000, 2008
Marzo/Abril/Mayo	
Niño moderado o fuerte	1983, 1987, 1992
Niño débil	1998, 2010
Años neutros	1981, 1882, 1984, 1986, 1988, 1990, 1991, 1993, 1994, 1995, 1996, 1997, 2001, 2002, 2003, 2204, 2005, 2006, 2007, 2009
Niña débil	1985, 1989, 1999, 2000
Niña moderada o fuerte	2008,
Junio/Julio/Agosto	
Niño moderado o fuerte	1987, 1997
Niño débil	1982, 1991, 2002, 2004, 2009
Años neutros	1981, 1983, 1984, 1986, 1989, 1990, 1992, 1993, 1994, 1995, 1996, 2001, 2003, 2005, 2006, 2009, 2007, 2008,
Niña débil	1985, 1998, 1999, 200, 2010
Niña moderada o fuerte	1988,
Septiembre/Octubre/Noviembre	
Niño moderado o fuerte	1982, 1987, 1997, 2002, 2009
Niño débil	1986, 1991, 1994, 2004, 2006
Años neutros	1981, 1985, 1989, 1990, 1992, 1993, 1996, 2001, 2003, 2005, 2008
Niña débil	1983, 1984, 1995, 2000
Niña moderada o fuerte	1988, 1998, 1999, 2007, 2010

Figura 7. Selección de años Niños, Niñas y Neutros según las temperaturas del ONI. Tomado de (Bautista-Cabezas, 2015)

Estos eventos clasificados por intensidad, han ocurrido en el Ecuador, lo que se muestra en las figuras 5 y 6 son basadas en el ONI ya que siempre se toma como referencia y se lo complementa con los estudios realizados en la región a la cual pertenece el Ecuador, que es la región niño 1+2; para lo cual en la figura 7 se muestra una comparación de ocurrencia de eventos en la región niño 3.4, región donde se calcula el ONI, y la región 1+2 que la región en la que forma parte el Ecuador que es el caso de estudio en este trabajo.

Años	Pacífico Central			Pacífico Oriental		
	El Niño 3.4	La Niña 3.4	Neutral 3.4	El Niño 1+2	La Niña 1+2	Neutral 1+2
1990			X			X
1991	X oct-dic					X
1992	X ene-jun			X feb-jun		
1993				X feb-jun		X
1994	X oct-dic				X mar-ago	
1995	X ene-feb	X sep-dic			X abr-ago	
1996		X ene-mar			X abr-dic	
1997	X may-dic			X mar-dic	X ene	
1998	X ene-may	X jun-dic		X ene-ago		
1999		X ene-dic			X abr-dic	
2000		X ene-dic			X ene-di	
2001		X ene-feb			X ene-dic	
2002	X ene.Ju-di				X ene	X
2003	X ene-mar				X mar-ago	
2004	X jul-dic					X
2005			X			X
2006	X ago-dic			X ago-dic		
2007	X ene	X ago-dic			X mar-dic	

Figura 8. Cronología de los episodios cálidos, fríos y neutros en las regiones Niño 3.4 y 1+2 en el océano Pacífico central y oriental.

Tomado de (Padilla, 2010)

El fenómeno de El niño y La niña ocurren en diferentes periodos de tiempos, no tienen ciclos fijos y por esta razón son monitoreados de manera constante, las regiones en donde alto índice de ocurrencia en en la región niño 3.4 (El norte de América) y la región niño 1+2 (incluye a Ecuador y Perú), en estas dos regiones mensualmente se realizan investigaciones en las cuales se puede visualizar posibilidades de ocurrencia de estos eventos, en los cuales están involucrados los factores que influyen en estos casos, los cuales serán explicadas a continuación.

Características determinantes del mar en los eventos oceanográficos extremos ENOS:

Existen 2 características principales del mar que aportan con información determinante en la ocurrencia de eventos extremos como lo es el ciclo ENOS; las mismas que son: La salinidad y la temperatura, a continuación se detallará las definiciones y datos importantes sobre cada una de ellas ya que aportan de gran manera en este trabajo.

Temperatura

La temperatura es una medida del calor o energía térmica de las partículas en una sustancia. (Hernans-Killam & Daou, 2001)

En el contexto de este trabajo la definición de temperatura es importante pero enfocada al mar y océanos debido a que son factores fundamentales dentro del tema en desarrollo.

Temperatura del océano

El principal aporte calorífico que tiene el agua del mar está representado por las radiaciones energéticas que le llegan del Sol. Su calor específico tiene un valor elevado en comparación con el calor específico de las demás sustancias existentes en la superficie del planeta; esto confiere al mar una extraordinaria capacidad para almacenar calor y por esta propiedad puede actuar como un gigantesco moderador del clima. Se entiende por calor específico, en general, la cantidad de calor necesario para aumentar en un grado centígrado la temperatura de un gramo de agua. (CIFUENTES-LEMUS, TORRES-GARCÍA, & FRÍAS-M, 1995)

Para mayor comprensión sobre la temperatura del océano primero se debe conocer que existe una división según su estructura térmica la cual está compuesta de las siguientes 3 capas: Superficial, limítrofe o también conocida como termoclina y las profundidades del océano. (Bergman, 2011)

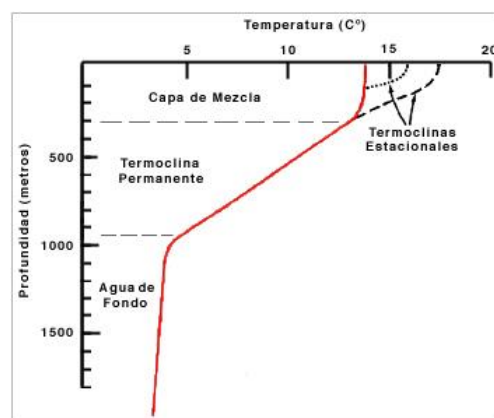


Figura 9. Estructura térmica del océano. Tomado de (CURSO GEOGRAFÍA DEL MAR, s.f.)

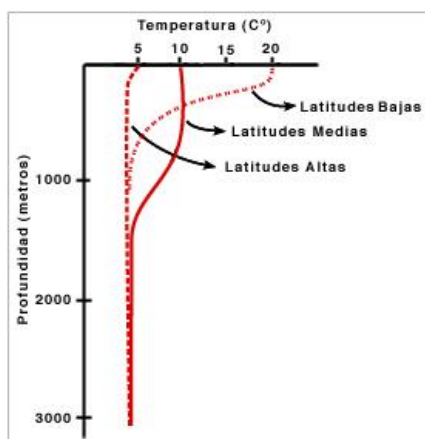


Figura 10. Estructura térmica del océano. Tomado de (CURSO GEOGRAFÍA DEL MAR, s.f.)

Como se puede observar en las Figuras 9 y 10 la temperatura del océano varía según la profundidad del mismo, a mayor profundidad menor temperatura y viceversa a menor profundidad mayor temperatura.

Temperatura superficial del mar (TSM)

La temperatura superficial del mar es el nivel térmico que se encuentra en la columna superficial del mar (0-200m) funciona como una interface en el intercambio de calor entre la atmósfera y el océano, es decir, es una característica física importante que influye en la transferencia del vapor de agua y de los gases entre el océano y la atmósfera. (Junta de Andalucía, 2009)

En Ecuador existen cuatro tipos de masas de aguas superficiales frente a la costa ecuatoriana según una investigación realizada por David Enfield (1976); estos tipos de masas de aguas encontradas se muestran en la tabla 1, detallando la temperatura y salinidad en condiciones normales para cada una debido a que dichas características del agua son muy importantes para el desarrollo de este trabajo.

<u>Tipo de masa de agua</u>	<u>Temperatura</u>	<u>Salinidad</u>
Aguas Subtropicales Superficiales (ASS)	>25°C	35 ups
Aguas Tropicales Superficiales (ATS)	>25°C	<33.0 ups
Aguas Ecuatoriales Superficiales (AES)	21°C a 25°C	33.0 a 34,5 ups
Aguas Costeras de Humboldt (ACH)	<21°C	>34,9 ups

Tabla 1. Tipos de masas de aguas superficiales frente a la costa ecuatoriana. Tomado de (Enfield, 1976)

Estas masas de agua superficiales son las que se usan para toma de muestras en el estudio de la temperatura superficial del mar y por consiguiente estos son los valores que se toman para detectar eventos extremos; de los ya antes mencionados; debido a que cada una de estas masas de aguas superficiales traen consigo características y organismos propios debido al rango de temperatura y salinidad que las mismas soportan.

Salinidad

Es el contenido de sales disueltas ya sea en el agua o en el suelo; para el caso de estudio de este trabajo se detallará sobre la salinidad en el agua.

La salinidad es considerada una propiedad química del agua la cual según el libro EL OCÉANO Y SUS RECURSOS II. LAS CIENCIAS DEL MAR: OCEANOGRAFÍA GEOLÓGICA Y OCEANOGRAFÍA QUÍMICA” (1997) dice que dicha propiedad resulta de la combinación de las diferentes sales que se encuentran disueltas en el agua oceánica, siendo las principales los cloruros, carbonatos y sulfatos.

En la tabla 1 se muestra el porcentaje de composición de cada elemento que aporta sales a los océanos.

Sales disueltas en las aguas de mar(átomos):
55.3 % Cloro
30.8 % Sodio
3.7 % Magnesio
2.6 % Sulfuro
1.2 % Calcio
1.1 % Potasio

Tabla 2. Sales disueltas en las aguas de mar. Tomado de (Bergman, Salinidad, Sales Disueltas, Medición de la Salinidad, 2001)

La concentración de sal en el agua del mar, denominada salinidad, es de aproximadamente 35 partes por mil, o lo que es lo mismo, un 3,5% del

peso del agua del mar proviene de las sales disueltas. (Zamorano, 2018). La medición de la salinidad es expresada en diferentes unidades; tal como ya se ha mencionado y expuesto una de las unidades de medida es ppm; pero existe otra que su uso es muy común esa unidad de medida es UPS (Unidades Prácticas de Salinidad), en este trabajo se usará dicha medida en los datos que serán procesados en su desarrollo, por dicha razón es importante su mención. Por lo tanto, la salinidad del agua de los océanos es de aproximadamente 35 UPS, este peso puede variar entre un océano y otro. (Bergman, 2001)

La salinidad varía en dirección tanto horizontal como vertical y aun en un mismo punto puede sufrir variaciones en las diferentes estaciones del año. Los factores que hacen cambiar la salinidad son, en primer lugar, la temperatura ya que si es elevada provoca una evaporación intensa y por lo tanto un incremento de salinidad resultante de la concentración de sales; en segundo lugar, los aportes de agua dulce, que por dilución, disminuye la salinidad. (LEMUS, GARCÍA, & M, 1997)

Como uno de los factores para variaciones en la salinidad es la temperatura, a continuación, en la Tabla 4 se muestra la relación entre la temperatura y la salinidad.

<i>Profundidad en metros</i>	<i>Temperatura 0° C</i>	<i>Salinidad</i>
0	26.44	37.45
50	18.21	36.02
100	13.44	35.34
500	9.46	35.11
1 000	6.17	34.90
1 500	5.25	34.05

Tabla 4. Relación Temperatura / Salinidad. Tomada de: (LEMUS, GARCÍA, & M, 1997)

Las variaciones generales de la salinidad están zonificadas del Ecuador a los polos. Los valores son bajos en el Ecuador, más altos en las regiones subtropicales y latitudes medias y bajos en las regiones polares. Los

principales procesos responsables de esta distribución son: la evaporación, la precipitación y, la mezcla. (Curso Geografía del Mar, s.f.)

En Ecuador existen valores bajos de salinidad debido a la alta evaporación que existe en su territorio, lo que provoca abundancia de precipitaciones (lluvias), también la gran cantidad de ríos que el país posee y sus desembocaduras en el océano son el factor principal para que las mezclas de aguas ecuatorianas tengan bajos niveles de sal en su composición; caso contrario si no tuviera alto índice de precipitaciones sus niveles de sal serían elevados tal y como sucede en otros lugares.

La salinidad se está se encuentra distribuida por zonas, las cuales pueden recibir diferentes corrientes que causan una alteración a esta distribución. Con las corrientes existentes en determinadas zonas del mar se podría determinar diferentes anomalías que podrían traer consigo eventos extremos oceanográficos los cuales a partir de estos datos acompañado de otros factores y con ello la presencia de algunos organismos marítimos indicadores se podría determinar su ocurrencia por lo cual se podría decir que la salinidad puede determinar los organismos que pueden vivir en el ciertas zonas del agua y por lo cual se convierte en un importante factor y característica en el medio ambiente.

Comportamiento de organismos marinos frente a la temperatura y la salinidad

Los hábitats de los organismos marinos tienen características muy acentuadas e importantes tales como la temperatura y salinidad del agua, las cuales como ya ha sido descrito anteriormente son influenciadas por diferentes factores los cuales hacen que los valores de dichas características sean variables; pero los organismos, *¿resisten todo rango de temperatura y salinidad del agua?* Los organismos también tienen rangos de adaptación y no todos resisten excesivos valores de temperatura o salinidad, ellos se mantienen en donde su naturaleza les permita y logren resistir y sobrevivir; es por esto que cuando existen variaciones en estas características del agua traen consigo a especies que son característicos de las mismas. Existe una

clasificación de las especies según su adaptación a la temperatura y salinidad.

Clasificación en grupos de adaptación a cambios de temperatura y salinidad

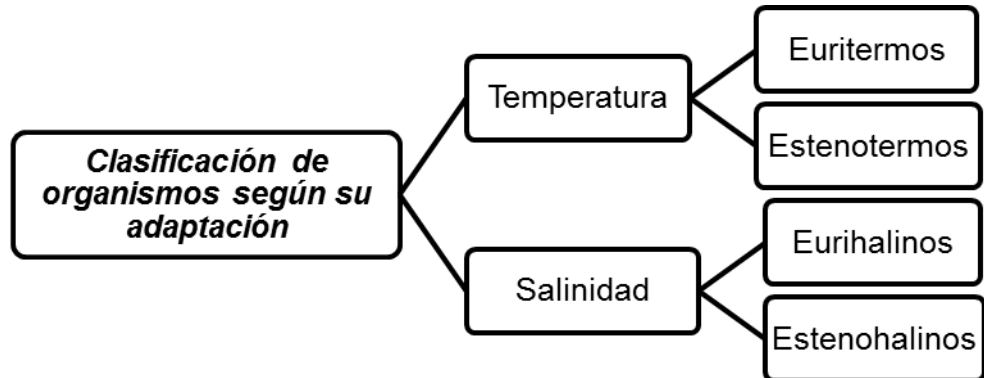


Gráfico 3. Clasificación de organismos según su adaptación. Autora

Clasificación de grupos de adaptación a cambios de Temperatura

Según Juan A. Alberto (2009) las definiciones para cada uno de ellos son las siguientes:

Euritermos: Organismos que soportan cambios importantes en las condiciones térmicas de su ambiente.

Estenotermos: Organismos que no soportan cambios importantes en las condiciones térmicas de su ambiente.

Clasificación de grupos de adaptación a cambios de salinidad

Según conceptos definidos en la publicación de Rafael Chávez Gómez, Arturo Rocha Ramírez y Alejandro Ramírez Rojas (2005) dicen que:

Eurihalinos: Se denomina así a los organismos que se adaptan a una variación amplia de salinidades.

Estenohalinos: Que no tolera grandes variaciones de salinidad.

Es importante mencionar estos grupos debido a que los organismos marinos de los cuales se habla en este trabajo (organismos zooplanctónicos) algunos están dentro de uno de estos 2 grupos; y por lo cual se muestra una

acentuada abundancia en ciertas condiciones, de vaciación de temperatura o salinidad; trayendo consigo detalles de las corrientes que han ingresado y de los cambios que se han dado, que como ya se sabe están muy asociados a la ocurrencia de eventos anómalos.

En Ecuador existen varias instituciones interesadas y dedicadas al tema de los fenómenos oceanográficos, el *INOCAR* es el pionero de este grupo pues es la primera institución en realizar cruceros de investigación para estudios oceanográficos; como dice su propio nombre es Instituto Oceanográfico de la Armada, es una institución dedicada a eventos oceanográficos en el país y el presente documento aborda ese tema; por lo cual es importante mencionarlo. Pero también es importante mencionar a otra importante institución en la rama de los organismos zooplanctónicos debido a que mediante este trabajo se quiere comprobar que el comportamiento de estos organismos en los ecosistemas marinos cumplen un destacado papel como organismos bioindicadores; por medio de datos de las investigaciones realizadas por el *INP* institución pública que tiene como objetivo: “Generar información y conocimiento científico-tecnológico para el aprovechamiento racional de los recursos hidrobiológicos y sus ecosistemas, proporcionando medidas de manejo y conservación a las autoridades competentes, contribuyendo al desarrollo del sector pesquero y acuicultor” (Instituto Nacional de Pesca, s.f.), ambas instituciones tienen como sede de sus instalaciones la ciudad de Guayaquil y ambas aportan con conocimientos científicos e informativos para la sociedad, el Ecuador y el mundo.

Técnicas de minería de datos

La minería de datos consiste en diversos conjuntos de procesos analíticos para explorar grandes conjuntos de datos. Los objetivos son diversos: el descubrir pautas de comportamiento o interrelaciones sistemáticas entre variables, generalmente de bases de datos empresariales, construir modelos predictivos, y en general extraer información no evidente utilizando métodos computacionales intensivos. El éxito en la aplicación de la minería de datos depende de dos factores: el planteamiento claro del problema y de los objetivos, y la disponibilidad de datos adecuados. La

calidad y fiabilidad de la información es importante, pues numerosas técnicas estadísticas son muy sensibles a la presencia de datos anormales o no representativos. En principio la minería de datos se aplica sobre bases de datos que se han obtenido sin ningún diseño muestral. (Caridad-Ocerin, 2011)

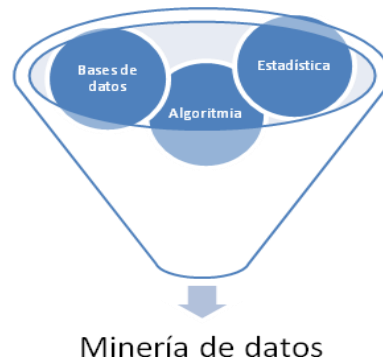


Figura 11. Minería de datos. Tomado de (Krall, 2006)

La minería de datos está basada en el uso de diferentes procesos y métodos inteligentes, los cuales ayudan a encontrar los patrones de los datos procesados, los algoritmos cumplen un rol fundamental en estos procesos, algunos de ellos son: arboles de decisión, redes neuronales, regresión lineal, reglas de asociación, entre otros; el conocimiento y análisis de patrones llevan a un conocimiento de los datos procesados, lo cual sirve para toma de decisiones y análisis relacionados con el caso de estudio o conocimiento hallado. La minería de datos se caracteriza por el procesamiento de grandes volúmenes de datos los cuales son extraídos de bases de datos, almacenes de datos (data warehouse) y combina los datos, la estadística y los algoritmos, tal y como se muestra en la figura 11.

La minería de datos está dividida en dos grupos según su tarea a realizar, dependiendo de las mismas se utiliza las técnicas adecuadas para cada una tal como se muestra en la figura 10.

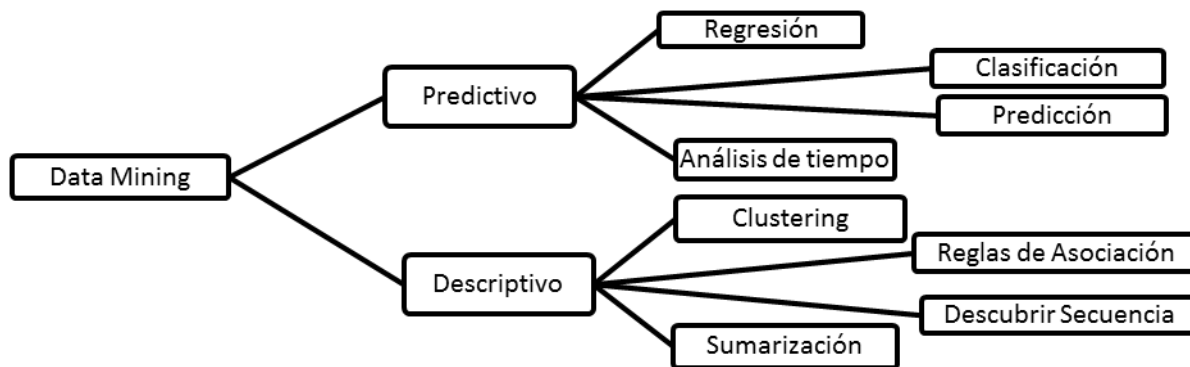


Figura 12. Clasificación en tipo de tareas de Minería de Datos. Tomado de Presentación elaborada por Odelmar Rodríguez con el tema: “Conceptos Básicos de Minería de Datos” (Rodríguez, 2008)

Tomando la clasificación mostrada en la figura 12 se puede detallar la definición para cada una de las tareas de minería de datos según lo mencionado en la presentación de Odelmar Rodríguez sobre Conceptos Básicos de Minería de datos (2008):

- *Predictiva*: Buscar patrones humano-interpretables que describen los datos
- *Descriptiva*: Utiliza algunas de las variables para predecir los valores futuros desconocidos de la misma variable o bien de otras variables.

Sabiendo que tipo de tarea será aplicada se analiza con cuál de estos métodos y técnicas de minería de datos se puede llegar al conocimiento de los datos y que también es un paso en el proceso KDD que consiste en la aplicación de análisis de datos y algoritmos de descubrimiento que, bajo limitaciones de eficiencia computacional aceptables, produce una enumeración particular de patrones sobre los datos. Teniendo en cuenta que el espacio de los patrones es a menudo infinito, y la enumeración de patrones implica alguna forma de búsqueda en este espacio. (Flores-Lagla, Cadena-Moreano, Quinatoa-Arequipa, & Quishpe, 2019)

Descubrimiento de Conocimiento a partir de Bases de Datos (KDD)

Es una Metodología minería de datos en la cual como ya antes ha sido mencionado la minería es parte del proceso que mediante algoritmos, técnicas y procedimientos se obtienen principalmente patrones de la información de bases de datos extensas e históricas. El proceso KDD tiene un modelo que cuenta con varios pasos ordenado para su ejecución, las cuales se muestran en la figura 8 a continuación:

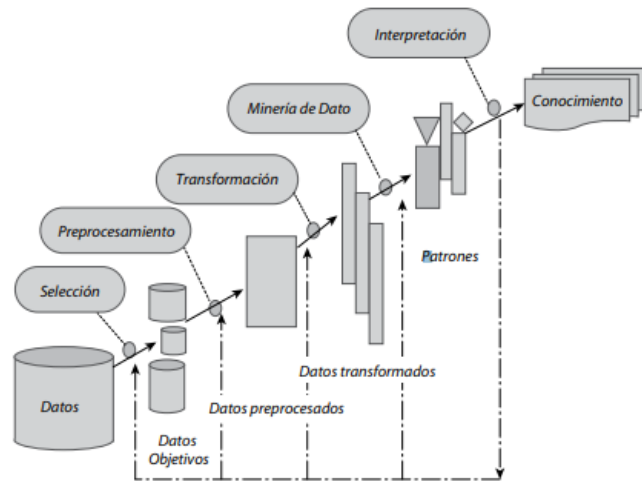


Figura 13. Proceso de Metodología KDD. Tomado de (Timarán, Hernández, Caicedo, Hidalgo, & Alvarado, 2016)

Tal y como se puede ver en la figura 13 los pasos de esta metodología son los siguientes: *selección, pre procesamiento y limpieza, transformación/reducción, minería de datos, interpretación/evaluación*, al cumplir con este proceso se llega objetivo planteado que es el conocimiento de los datos, mediante un modelo que permite el proceso de los datos y la obtención de patrones o de elementos claves para estudio o toma de decisiones referente al tema en cuestión.

Existen varias metodologías de minería de datos las cuales van a ser nombradas a continuación en la figura 14, en la cual se realizará una comparación entre las etapas de cada una.

KDD	SEMMA	CRISP – DM
Pre – KDD	-----	Comprensión del Negocio
Selección	Muestreo	Comprensión de los Datos
Preprocesamiento	Exploración	Preparación de los Datos
Transformación	Modificación	Modelado
Minería de Datos	Modelado	Evaluación
Interpretación/Evaluación	Evaluación	Implementación
Post KDD		

Figura 14. Cuadro comparativo metodologías de Minería de datos.
Tomado de (GUTIÉRREZ-PACHERRES, 2017)

Como se muestra en la figura 14, la metodología que cuenta con un proceso de etapa más completo es la *metodología KDD*, y pues se puede decir que las otras metodologías expuestas en el cuadro parten del proceso de KDD; por lo cual esa es la razón por la que este trabajo se utilizará dicha metodología debido a que cuenta con un conjunto de etapas completas en las cuales su objetivo principal es la detección de patrones y de los cuales obtener un nuevo conocimiento; lo que para esta investigación aporta y forma parte de su desarrollo para el cumplimiento de sus objetivos.

Patrones

En minería de datos los patrones sirven para interpretar los datos, un patrón es algo que ocurre repetidamente o que cumple reglas, tendencias o características específicas, entonces muestra el comportamiento del fenómeno, suceso u objeto que se está estudiando, por ejemplo en el caso del presente trabajo los patrones de distribución que se detectarán nos darán resultados del comportamiento de los organismos zooplanctónicos en su ecosistema y frente a las características y condiciones del mismo que serán parte del estudio, con los cuales se espera obtener los patrones analizarlos de manera adecuada y con conocimientos necesarios para poder así generar el modelo propuesto.

Modelo de minería de datos

Un modelo es una forma de aplicar un tratamiento a una cantidad masiva de datos para extraer información de ellos. (Krall, 2006)

El modelo de minería de datos es el producto de el uso de algoritmos, los cuales hacen que este sea mas que solo un dato, sino que se convierta en un conjunto de datos que sirven para predicciones y deducciones, como por ejemplo en este trabajo lo que se está buscando es generar un modelo predictivo detectando patrones de distribución de los organismos zooplanctónicos, el cual se espera que proporcione predicciones de eventos extremos de la variabilidad climática mediante el modelo que se genere.

Definir y crear el modelo de minería de datos es todo un proceso, el cual tiene varios pasos que se deben seguir para que pueda ser realizado de la mejor manera, estos pasos recomendados se los enumera a continuación en el grafico 2:

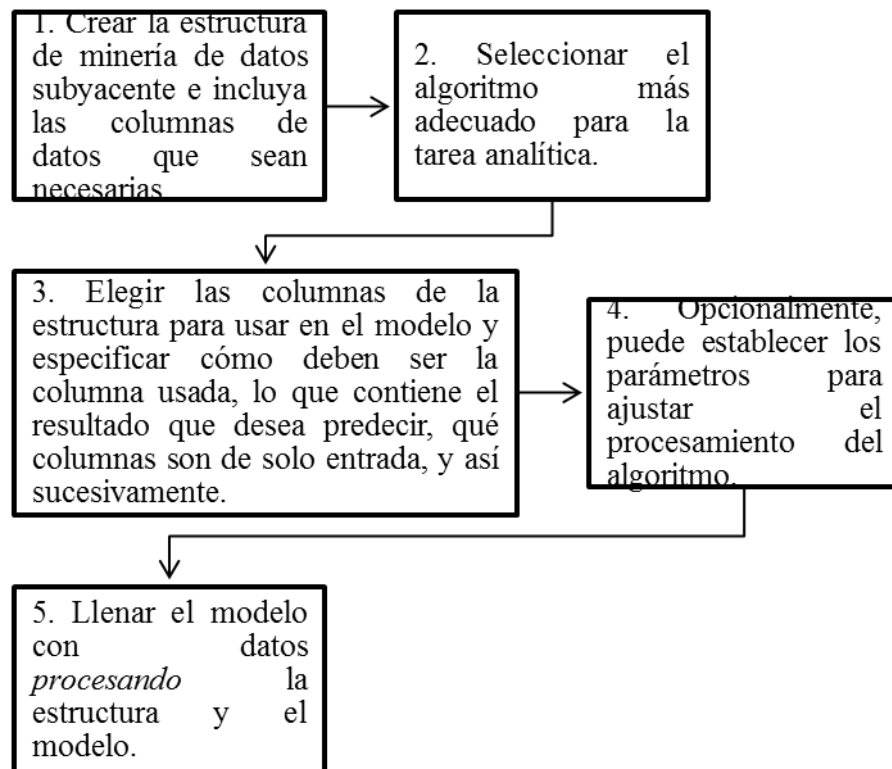


Gráfico 2. Pasos para definir el modelo de minería de datos. Tomado de (Microsoft, 2018)

Siguiendo estos pasos se define el modelo que será presentado y el cual procesará los datos para poder tener información clave sobre la investigación o tema planteado, y así llegar al conocimiento que es el final del proceso KDD que es la metodología aplicada en este trabajo.

Algoritmos de minería de datos

Un algoritmo es un conjunto detallado y lógico de pasos para alcanzar un objetivo o resolver un problema. (Manzano, Montesano, & Zúñiga, 2017). En minería de datos los algoritmos son usados para llegar al modelo predictivo, y para poder detectar los patrones que son parte del modelo. En el gráfico 3 se detalla las características que debe cumplir un algoritmo para que

Finito	Dentro de la secuencia de pasos para realizar la tarea, debe tener una situación o condición que lo detenga; de lo contrario, se pueden dar ciclos infinitos que impidan llegar a un término.
Preciso	Un algoritmo no debe dar lugar a criterios.
Obtener el mismo resultado	En cualquier circunstancia, si se atienden en forma estricta los pasos del algoritmo, siempre se debe llegar a un mismo resultado. Ejemplos: obtener el máximo común divisor de dos números enteros positivos, armar un modelo a escala, resolver una ecuación, etcétera.

cumpla su propósito.

Gráfico 3. Características de un algoritmo. Tomado de (Manzano, Montesano, & Zúñiga, 2017)

Cumpliendo con las características de un algoritmo, los cuales en la minería de datos ayudan a la obtención de los patrones y la información que estos brindan, los algoritmos en minería de datos según una publicación de Yuniet Rodríguez Suárez y Anolandy Díaz Amador (2009) tienen la siguiente clasificación:

Supervisados o predictivos: predicen el valor de un atributo de un conjunto de datos, conocidos otros atributos. A partir de datos cuya etiqueta se conoce, se induce una relación entre dicha etiqueta y otra serie de atributos. Esas relaciones sirven para realizar la predicción de datos cuya etiqueta es desconocida.

No supervisados o del descubrimiento del conocimiento: con estos algoritmos se descubren patrones y tendencias en los datos actuales. El

descubrimiento de esa información sirve para llevar a cabo acciones y obtener un beneficio de ellas.

Según esta clasificación, el tipo de algoritmos que se utilizarán en este trabajo serán los *supervisados o predictivos*, debido a que tal y como lo dice el nombre se generará un modelo para la predicción de eventos extremos por lo cual a partir de una serie de característica o atributos se realizará la predicción.

Las normativas, reglamentos, decretos en los **MARCO LEGAL** cuales se involucra los datos e investigaciones en este trabajo son las siguientes:

1. *Creación de INP y designación de funciones*

Según lo mencionado en los registros oficiales de la institución (Instituto Nacional de Pesca, s.f.) se indica que:

El Instituto Nacional de Pesca (INP) es una entidad de derecho público creada el 5 de diciembre de 1960, con personería jurídica, patrimonio y recursos propios. Está adscrito Ministerio de Acuacultura y Pesca (MAP) y tiene su domicilio en la ciudad de Guayaquil.

Del Acuerdo Básico entre el Gobierno del Ecuador y el Fondo Especial de las Naciones Unidas, para la ejecución de Proyectos de Desarrollo Económico (10 de noviembre de 1959), y a pedido del Ministerio de Fomento, mediante Decreto No. 582-a, del 5 de diciembre de 1960, se crea el Instituto Nacional de Pesca, cuya existencia legal descansa el 5 de enero de 1961, según el Registro Oficial No. 105, e inicia sus actividades el 7 de mayo del mismo año. Por recomendación de la FAO (Food and Agriculture Organization), el Instituto Nacional de Pesca (INP) pasa a ser un organismo especializado dedicado a la investigación biológica, tecnológica y económica, tendientes a la ordenación y desarrollo de las pesquerías (Decreto Ejecutivo No. 1321 del 18 de octubre de 1966).

El INP cumpliendo sus funciones decretadas en su creación, constantemente realiza investigaciones que aportan en la productividad del

país, y no solo eso, sino que dichos estudios dan un aporte científico en el resto del mundo.

El estudio de los organismos zooplanctónicos es parte de sus investigaciones biológicas, las cuales se realizan enfocadas para resultados oceanográficos o pesqueros; en los estudios oceanográficos se tiene al evento “El Niño/La Niña” que debido a la creación de un comité regional se realizan estudios y monitoreos mensuales en conjunto con otras instituciones miembros tal como se describe a continuación;

2. Creación de Comité ERFEN (Estudio Regional del Fenómeno de "El Niño")

En Registro Oficial N° 293, el día martes 27 de marzo del 2001, siendo Gustavo Noboa Bejarano Presidente Constitucional de la República del Ecuador y considerando:

Que la ocurrencia del Fenómeno del Niño, tiene muchas repercusiones en la población y causa pérdidas en los diferentes ámbitos productivos económicos del Ecuador, y las costas ecuatorianas situadas en el Pacífico Sur; dentro la Comisión Permanente del Pacífico Sur (CPPS) se crearon los comités generales para el Estudio Regional del Fenómeno de “El Niño”(ERFEN) ; como organismos científico-técnicos, que funcionan con la participación de las principales instituciones de investigación al nivel de cada país integrante de CPPS y con el soporte de otras organizaciones internacionales. (INOCAR)

Se decreta la Creación del ERFEN que consta de 5 artículos:

Art 1.- Creación y Sede. - En este artículo se define a la ciudad de Guayaquil como sede del comité; en donde se ejercerán todas las funciones con respecto a la ocurrencia del evento “El Niño/La Niña”

Art 2.- Integración: En este artículo se define que entidades e instituciones integrarán este comité regional; todas relacionadas con estudios oceanográficos y ambientales. Dentro de las instituciones nombradas en este artículo se encuentra el INP; la cual aporta con sus investigaciones

oceanográficas y de organismos marinos relacionados al tema de la ocurrencia de estos eventos.

El INP según la información tomada de (Instituto Nacional de Pesca, s.f.) indica que: como miembro del Comité Nacional ERFEN aporta con información pesquera y oceanográfica, mediante el monitoreo de los recursos pesqueros que puedan mostrar algún cambio en el comportamiento del océano, así como migraciones o desplazamientos de peces pelágicos pequeños, peces pelágicos grandes, camarón, merluza, entre otros.

Además, mediante su programa de variabilidad climática y cruceros, realiza monitoreos *in situ* en el océano, donde se toman parámetros físicos, químicos y biológicos del mar, con el fin de conocer los diferentes cambios oceanográficos causados por la estacionalidad y eventos extremos como el niño y la niña.

Se hace mención a esta institución debido a que los datos para la ejecución de este trabajo son parte de sus investigaciones, las cuales están basadas en las funciones adjudicadas en este comité.

Art 3.- Competencias: En este artículo se detallan las funciones que deben tener los miembros de este comité; los cuales deberán constantemente estar impulsando y promoviendo estudios referentes al evento; y no solo eso sino ser los informantes continuos y asesorar de manera oportuna al Gobierno Nacional con respecto a ocurrencia y estudios sobre desastres costeros.

Art 4.- Dirección y Coordinación: En este artículo se designan responsabilidades para la coordinación del comité regional creado.

Art 5.- En este artículo se detalla como función la asesoría del comité a la Dirección Nacional de Defensa Civil u otros organismos nacionales encargados de la elaboración de planes de prevención y contingencia ante la presencia de eventos naturales adversos.

Este decreto está vigente desde su creación en Quito el 16 de marzo del 2001; desde entonces las instituciones involucradas han cumplido con sus investigaciones y monitoreo del evento del niño; mantienen reuniones de

manera periódica de las cuales se generan boletines informativos del estado del océano ante el Evento “El Niño/La Niña” los cuales son parte de esta investigación.

3. *Plan del Buen vivir 2017-2021 (Toda una Vida)*

En la sección “Lineamientos territoriales para cohesión territorial con sustentabilidad ambiental y gestión de riesgos” el literal b.1 dice lo siguiente:

“Implementar procesos para la identificación, conocimiento, conservación y revalorización de los paisajes naturales y culturales, terrestres, acuáticos y marino-costeros, que aseguren su integridad, conectividad y funcionalidad, como condición básica para la generación de servicios ambientales esenciales para el desarrollo sostenible”.
(Secretaría Nacional de Planificación y Desarrollo, 2017)

Cumpliendo este literal la realización de este este trabajo destaca el ecosistema marino-costero como indicador de ocurrencia de eventos anómalos; lo cual es esencial para el desarrollo sostenible tal como lo menciona, debido a que con información y monitoreo de posibles fenómenos o desastres se mantiene la integridad del país y de sus habitantes, ya que se mantienen informados y prevenidos.

Los 3 numerales mencionados están relacionado con el estudio del ecosistema marino-costero; en los cuales están involucrados los organismos zooplanctónicos y los eventos extremos de la variabilidad climática (“El niño/La niña”), quienes son parte de este trabajo; siguiendo las normas del buen vivir y decretos mencionados en donde se involucra este trabajo el mismo cumple como un aporte en el tema ambiental de ocurrencia de eventos extremos.

CAPÍTULO III

METODOLOGÍA DE LA INVESTIGACIÓN

El presente trabajo tiene un enfoque *cuantitativo* debido a que la investigación parte de la hipótesis que los organismos zooplanctónicos son indicadores de ocurrencia de los eventos extremos de la variabilidad climática y tiene como objetivo obtener los patrones de distribución de dichos organismos los cuales mediante un análisis permitan llegar a conclusiones que puedan comprobar o descartar la hipótesis planteada.

Es importante mencionar la definición del enfoque cuantitativo y así poder indicar que esta la misma cumple con las características de este trabajo; entonces la definición del enfoque cuantitativo de la investigación nos dice que el mismo usa la recolección de datos para probar hipótesis, con base en la medición numérica y el análisis estadístico, para establecer patrones de comportamiento y probar teorías. (Hernández-Sampieri, Fernández-Collado, & Baptista-Lucio, 2005)

El enfoque cuantitativo cuenta con varios tipos de investigación que pueden ser usadas; para el desarrollo de este trabajo se usará la investigación *descriptiva* debido a que se quiere saber el comportamiento predominante de los organismos en estudio que son parte de la hipótesis.

Se define que este trabajo usa la investigación descriptiva debido a que el concepto de la misma dice que su objetivo consiste en llegar a conocer las situaciones, costumbres y actitudes predominantes a través de la descripción exacta de las actividades, objetos, procesos y personas. Su meta no se limita a la recolección de datos, sino a la predicción e identificación de las relaciones que existen entre dos o más variables. Los investigadores no son meros tabuladores, sino que recogen los datos sobre la base de una hipótesis o teoría, exponen y resumen la información de manera cuidadosa y luego

analizan minuciosamente los resultados, a fin de extraer generalizaciones significativas que contribuyan al conocimiento. (Dalen & Meyer., 2006)

Mediante el uso del *método deductivo* se podrá llegar a las conclusiones adecuadas, que por medio del principio general se pueda establecer los patrones de distribución de los organismos zooplanctónicos y con dichos patrones generar el modelo predictivo propuesto.

La definición del método deductivo dice que es una forma de razonamiento que parte de una verdad universal para obtener conclusiones particulares. En la investigación científica, este método tiene una doble función encubrir consecuencias desconocidos de principios conocidos. (López-Cano, 1986)

Con el enfoque, métodos y técnicas de la investigación que se usarán se podrá lograr un trabajo estructurado, entendible y con los resultados deseados, debido a que el seguimiento de cada uno aporta a diferentes partes del desarrollo y la investigación permitiendo llegar al cumplimiento de los objetivos planteados de una manera adecuada.

La metodología de **METODOLOGÍA DE MINERÍA DE DATOS** minería de datos a usar en esta propuesta es KDD (Knowledge Discovery in Databases) que significa Descubrimiento de Conocimiento en Base de Datos según la publicación “El proceso de descubrimiento en base de datos” escrita por varios autores (2016) esta metodología es un proceso que consta de cinco fases que son las siguientes:

- *Selección:* En la etapa de selección, una vez identificado el conocimiento relevante y prioritario y definidas las metas del proceso KDD, desde el punto de vista del usuario final, se crea un conjunto de datos objetivo, seleccionando todo el conjunto de datos o una muestra representativa de este, sobre el cual se realiza el proceso de descubrimiento. La selección de los datos varía de acuerdo con los objetivos del negocio. (Timarán, Hernández, Caicedo, Hidalgo, & Alvarado, 2016)

Esta fase está dada por la obtención de la base de datos de la distribución y abundancia de los organismos zooplanctónicos en las costas ecuatorianas, datos proporcionados por el INP, encargado de realizar estudios relacionados con dichos organismos los cuales son parte fundamental de este trabajo.

La base de datos proporcionada posee información que ha sido generada de los trabajos de campo realizados mensualmente por dicha institución, los cuales se extraen de muestras de aguas tomadas en las costas ecuatorianas.

- *Pre procesamiento y limpieza:* En esta etapa se determina la confiabilidad de la información, es decir, realizar tareas que garanticen la utilidad de los datos. Para esto se hace la limpieza de datos (tratamiento de datos perdidos o remover valores atípicos). Esto implica eliminar variables o atributos con datos faltantes o eliminar información no útil para este tipo de tareas como el texto (aunque puede utilizarse para hacer Minería de Texto, que es otro asunto). (Landa, 2016)

El análisis de la base de datos a utilizar, revisión de cada campo de la tabla, para poder realizar una limpieza de campos incompletos o vacíos, debido a que dichos campos no aportarían en el resultado exacto para el modelo predictivo que se generará.

Como el periodo de estudio es extenso, pueden existir ausencias de datos en ciertos registros los cuales deberán ser revisados para cumplir con el objetivo del modelo de la manera más exacta.

- *Transformación/reducción:* En la etapa de transformación/reducción de datos, se buscan características útiles para representar los datos dependiendo de la meta del proceso. Se utilizan métodos de reducción de dimensiones o de transformación para disminuir el número efectivo de variables bajo consideración o para encontrar representaciones invariantes de los datos. (Fayyad, Piatetsky-Shapiro, & Smyth, 1996)

Una vez que los datos obtenidos de la base son correctos y servirán para la minería se verifica que el formato de la base de datos sea compatible para realizar el proceso.

Los datos que se poseen deben ser los registros de la data que pueden ayudar a determinar aspectos importantes en el presente trabajo para el diseño de la vista minable en donde se realiza el reconocimiento de los diferentes casos de estudios para poder continuar el proceso de minería.

- *Minería de Datos*: La definición formal de la etapa de data Mining o minería de datos sería, la extracción no trivial de información implícita, previamente desconocida y potencialmente útil a partir de datos. Otra manera de definirlo podría ser: la exploración y el análisis -por medios automáticos o semiautomáticos- de grandes cantidades de datos con el fin de descubrir patrones con significado. (Sinchiguano-Vizuet, 2017)

En esta etapa mediante algoritmos y técnicas se obtienen los patrones que se están buscando para poder generar el modelo predictivo propuesto.

Las técnicas y algoritmos de minería de datos están divididas según su tipo en *predictivas* y *descriptivas*; en la figura 15 se muestra el detalle de



esta clasificación con las técnicas y algoritmos aplicados para cada una.

Figura 15. Técnicas de minería de Datos. Tomado de (García-Herrero & Molina-López, 2012)

- *Interpretación y evaluación:* Interpretación de los patrones minados y generar el modelo predictivo planteado; luego de un análisis detallado de los resultados (patrones) obtenidos en el proceso de minería ejecutado en la fase anterior. Los resultados obtenidos serán interpretados con ayuda de las personas especializadas en el tema para que los mismos sean más exactos y eficaces.

El prototipado **METODOLOGÍA DE PROTIPADO RAPIDO** rápido se describe como un método basado en ordenador que pretende reducir el ciclo iterativo de desarrollo. Los prototipos iterativos desarrollados podrán ser rápidamente reemplazados o modificados según los informes de diversas procedencias, como experiencias previas de usuarios o de diseñadores veteranos, a medida que se evoluciona en el desarrollo de las tareas a realizar. (Cortés, 2000)

Este método será usado en este trabajo para la muestra de resultados y comprobación del modelo propuesto, por lo cual no se necesita un proceso extenso de desarrollo, debido a que el software a presentar será sencillo y cumplirá con especificaciones básicas, en los cuales los datos tomados como resultados serán presentados.

Etapas de la metodología de prototipado rápido

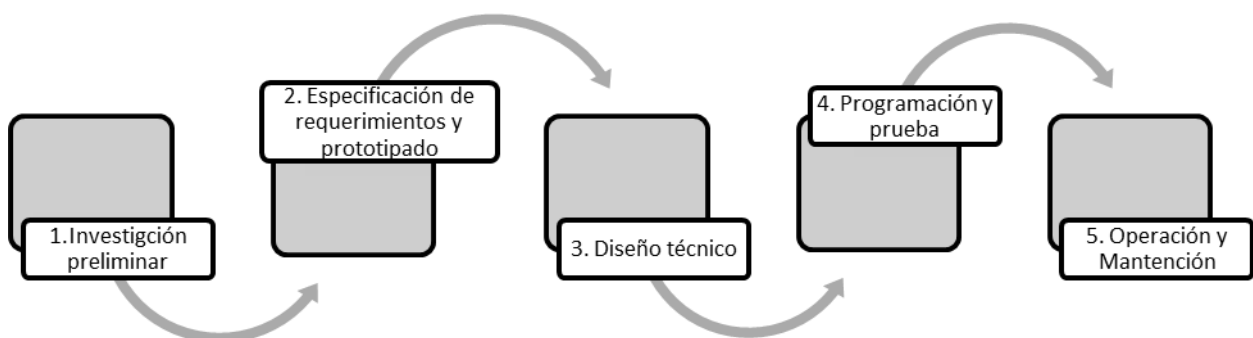


Figura 16. Etapas de la metodología de prototipado rápido. Tomado de (García & I., 2011)

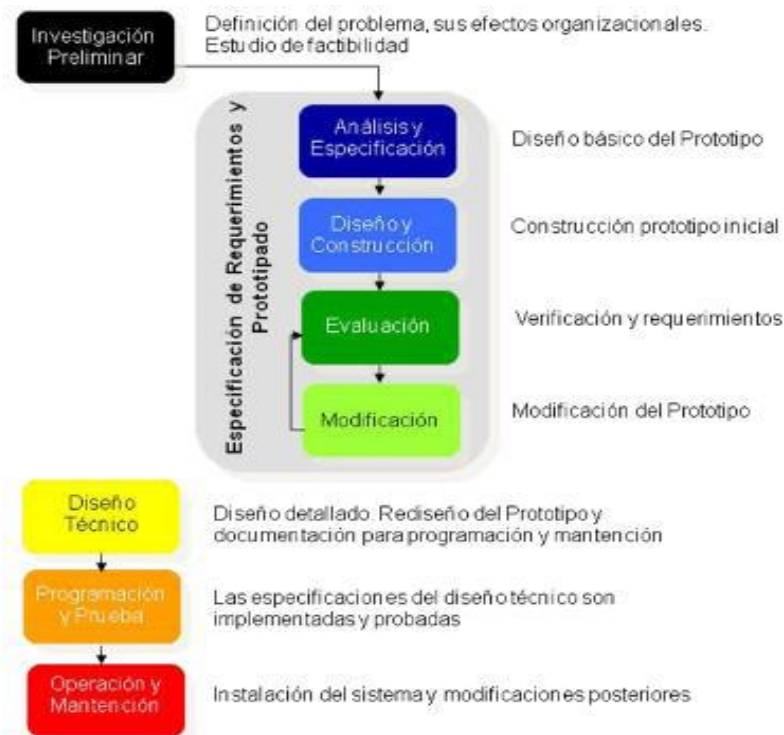


Figura 17. Detalle de las etapas del prototipado rápido. Tomado de (García & I., 2011)

El instrumento de recolección de datos utilizado es la *entrevista abierta*, debido a que es importante tener las explicaciones especializadas de las personas encargadas de los estudios relacionados con los organismos zooplanctónicos y la ocurrencia de fenómenos extremos de la variabilidad climática que para el caso de este trabajo es el ciclo ENOS. Las personas que fueron entrevistadas pertenecen al INP, debido a que es la institución que proporcionó los datos para el desarrollo de este trabajo.

Las entrevistas fueron realizadas a tres personas claves en el proceso y estudio del Ciclo ENOS y los organismos zooplanctónicos. Algunas de las preguntas se las realizaron

a los tres entrevistados y otras se derivaron de respuestas dadas y dependiendo del rol que cumple y la función que realiza en el proceso la persona a la cual se le realizaba la entrevista. Para mayor conocimiento y comprensión a continuación en la tabla 5 se mostraran los roles (cargos) de las tres personas que fueron entrevistadas.

ROL ENTREVISTADO	ENTREVISTA
INVESTIGADOR PESQUERO	E1
OCEANÓGRAFO	E2
DIRECTORA INP	E3

Tabla 5. Roles de entrevistados. Autora

Tal y como se puede ver en la tabla 5 los roles de cada una de las personas a las que se le realizó la entrevista cumple un rol importante en el proceso y estudio de la ocurrencia del ciclo ENOS, cada cual con su grado de experticia en su especialidad y función que cumple en este proceso.

<i>Preguntas realizadas en las tres entrevistas</i>
¿Cómo define Ud. el ciclo ENOS (El niño/oscilación sur)?
¿Qué trabajos e investigaciones realiza el INP con respecto a la ocurrencia de estos eventos?
¿Cuál es la frecuencia con la que se realizan los estudios relacionados al monitoreo del ciclo ENOS?
¿Considera que los organismos zooplanctónicos son indicadores de ocurrencia del ciclo ENOS? ¿Por qué?
Actualmente, ¿se realizan predicciones o solo se monitorea estos eventos?

Tabla 6. Preguntas de entrevista. Autora

Como ya fue mencionado anteriormente existieron preguntas en común en las entrevistas, las cuales nos permiten decir que el INP con

respecto al tema de eventos extremos y organismos zooplanctónicos tiene un programa permanente denominado: “Variabilidad Climática”, por medio del cual la institución ha creado estaciones fijas de trabajo (para toma de muestras, puntos específicos en las costas ecuatorianas), las cuales comprenden desde el norte hasta el sur, en las cuales mensualmente realizan el trabajo de campo, con el cual monitorean si están ingresando corrientes cálidas desde el norte o corrientes frías desde el sur. Para este programa se cumple un proceso en el cual tiene como actores las tres personas entrevistadas.

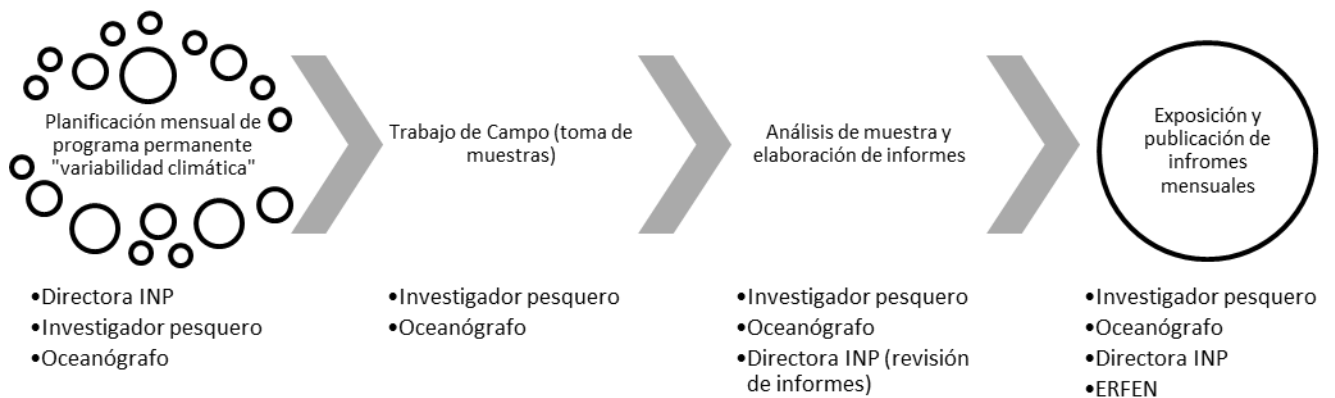


Gráfico 4. Proceso de Programa permanente del INP: “Variabilidad Climática”. Autora

Los entrevistados explicaron que los informes mensuales que realizan con respecto al programa relacionado al ciclo ENOS no solo son expuestos en el comité ERFEN sino que después de ahí se expone en el CPPS (Comité Permanente Pacífico Sur).

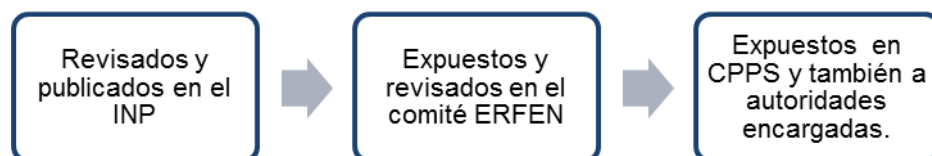


Gráfico 5. Proceso de publicación y exposición de informes de Programa permanente del INP: “Variabilidad Climática”. Autora

Como información clave en estas entrevistas todos coincidieron en que el desarrollo de un modelo predictivo para ocurrencia de los eventos del ciclo ENOS es de gran importancia y avance en este tema, debido a que en la actualidad no existe un modelo predictivo, el proceso que se realiza es netamente estadístico comparativo y en ocasiones se realizan proyecciones.

Debido a que el INP cumple como suministro de información para las autoridades un modelo predictivo ayudaría en el proceso con información más precisa y con carácter preventivo si fuera necesario.

Detalle de entrevistas

E1 - Investigador Pesquero

La entrevista tuvo como tema central el ciclo ENOS y el rol del investigador pesquero dentro de las investigaciones y trabajos realizados referentes al tema. Por lo que la entrevistada pudo indicar que se realiza un monitoreo constante en el cual su participación está presente en todas las fases del programa permanente de variabilidad climática.

La investigadora pesquera en esta entrevista destaca las variables que toma en cuenta para el análisis de estos estudios. Y los clasifica de la siguiente manera:

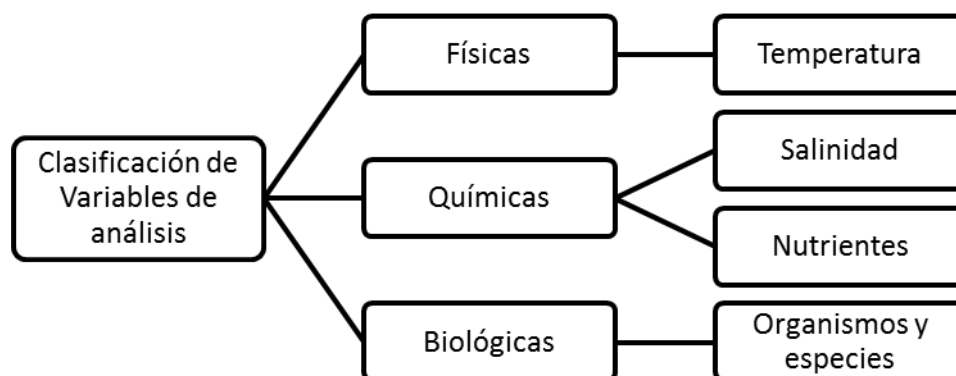


Gráfico 5. Clasificación de Variables de análisis. Autora

Teniendo en cuenta las variables que mencionó, una vez realizado el trabajo de campo, que para su caso es la toma de muestras, se analizan y muestra resultados con relación a las variables, ya que indican condiciones especiales o normales en el océano.

E2 – Oceanógrafo

La entrevista realizada con respecto a su rol dentro de los trabajos e investigaciones del ciclo ENOS, el entrevistado pudo mencionar que en sus trabajos lo que utiliza para el procesamiento de los datos son modelos matemáticos y estadísticos, con los cuales mensualmente realiza sus informes. Destacó que el comportamiento de los organismos zooplanctónicos traen consigo características especiales de las masas de agua presentes en el mar, y por lo cual si considera que son indicadores de ocurrencia de eventos extremos.

El entrevistado considera que los estudios realizados sobre este tema son para toma de decisiones, debido a que se trata de un fenómeno climático que puede llegar a ser devastador, y con oportuna información, monitoreo y estudio, se podría realizar contingencias, por parte de las entidades encargadas de este tipo de temas.

E3 – Directora INP

La entrevista estuvo enfocada respecto a la toma de decisiones e importancia del monitoreo constante del ciclo ENOS, la directora indicó que el ciclo ENOS tiene ventajas y desventajas, y debido a su gran impacto. Cumpliendo su rol es parte del proceso del programa permanente que tiene el INP denominado variabilidad climática, aportando con las investigaciones de la institución y siendo parte del ERFEN en el cual se exponen resultados de dichos trabajos.

Actualmente indicó que la institución no cuenta con un modelo predictivo para este programa de variabilidad climática, que todos son presentados mediante tendencias y herramientas estadísticas, las cuales muestran resultados del monitoreo mas no cumplen la función predictiva.

CAPÍTULO IV

PROPUESTA

En este capítulo se desarrollará el modelo predictivo planteado, mediante el uso de minería de datos y su metodología KDD, cumpliendo cada etapa y desarrollando cada proceso propuesto en las mismas. A continuación se detallará lo desarrollado en cada etapa de la metodología en la herramienta de minería de datos elegida según sus características.

SELECCIÓN DE LA HERRAMIENTA PARA IMPLEMENTACIÓN DE METODOLOGÍA KDD

La herramienta en la que se hará la implementación de la metodología KDD, es KNIME Analytics Platform, es un software de código abierto el cual cuenta con los algoritmos necesarios para realizar el modelo predictivo propuesto, a continuación se expondrá en una tabla comparativa sus características.

Tabla 7. Comparación de herramientas de minería de datos. Tomado de (Lara Alfredo, 2014)

Herramientas de minería de datos	tipo de software	plataforma	algoritmos	tipo de modelo
Weka	Libre	Todas las plataformas	Clustering, Regresión	Predictivo
Clementine	Libre	Windows, Linux	Red neuronal, GRI A priori, logística, QUEST, CHAID, KARMA	Predictivo
Knime	Libre	Windows, Linux, Mac OS	Algoritmos segmentación, árboles de	Predictivo

			decisión, redes neuronales, SVM	
IBM SPSS	Libre	Windows, Linux	Ecuaciones estructurales	Predictivo
RapidMiner	Libre	Windows, Linux	Clustering, arboles de decisión, redes neuronales	Predictivo

Nota: Como se puede ver en el cuadro las herramientas de minería de datos cuentan con características similares pero KNIME sobresale debido a su variedad de algoritmos con los que cuenta y su característica multiplataforma.

La última actualización (noviembre 2018) del cuadrante de Gartner ubica a la herramienta KNIME entre los líderes, lo cual hace que su reputación, aceptación y popularidad sean muy buenas y por lo cual hace de esta herramienta la opción idónea e indicada para el desarrollo del modelo predictivo.



Figura16.Diagramade Gartner Herramienta de minería de datos.

1. Selección

IMPLEMENTACIÓN METODOLOGÍA KDD

de Datos

Los datos seleccionados, fueron obtenidos de la base de datos del Programa de Variabilidad Climática realizado por el INP, en los cuales constan registros mensuales de los últimos diez años, por lo cual se posee una base de datos con gran volumen de datos, que provienen de las tomas de muestras asociadas a este estudio, y son almacenados en un archivo con formato xlsx (libro de Excel).

La base de datos consta de los siguientes campos:

<i>Columnas de la base de datos</i>	<i>Tipo de dato</i>	<i>Descripción</i>
FUENTE	String	Nombre de proyecto
ESTACIÓN	String	Lugar del muestreo
HORA DE COLECTA	Date time	Hora en la que fueron recolectadas las muestras
DIA	String	Día en la que fueron recolectadas las muestras
MES	String	Mes en la que fueron recolectadas las muestras
AÑO	Number (integer)	Año en la que fueron recolectadas las muestras
LATITUD	String	Posición geográfica
LONGITUD	String	Posición geográfica
PROF. CORR	String	Profundidad en la que fue tomada la colecta de muestra
TIPO DE ARRASTRE	String	Método de colecta de muestra
CLASE	String	Grupos de organismos
ORDEN	String	Familia a la que pertenece el organismo
GÉNERO	String	Nombre principal de especie- nombre científico
ESPECIE	String	Nombre del organismo
NOMBRE CIENTÍFICO DE LA ESPECIE	String	Nombre único para referirse al organismo

ESPECIE	Number(integer)	Número de organismos encontrados
Transformados (org.m⁻³)	Number (double)	Numero de organismos encontrados elevados a metros cúbicos
Responsable del análisis	String	Persona que realiza el análisis
Llenado por	String	Persona que ingresa información a la base de datos
Temperatura	Number (double)	Valor de la temperatura superficial del mar
Salinidad	Number (double)	Valor de la salinidad presente en el mar
Disco sechi	Number (Integer)	Valor de profundidad de transparencia.

Tabla 8. Diccionario de datos. Autora

En esta etapa se cargó el archivo de la base de datos en la herramienta, usando el nodo *Excel Reader (XLS)*.

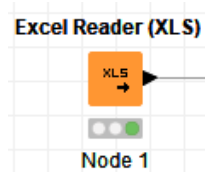


Figura 17. Nodo Excel Reader (XLS). Tomado de KNIME Analytics Platform

Con el uso de este nodo los datos quedan cargados en la herramienta mostrándose la primera vista preliminar tal y como se obtuvo el archivo de la

Output table - 2:1 - Excel Reader (XLS)

File Hilite Navigation View

Row ID	S	profundi...	S	TIPO DE...	S	CLASE	S	ORDEN	S	GÉNERO	S	ESPECIE	S	Nombre ci...	I	Especie	D	Transfor...	S	Respons...	S	Llenado ...	D	Temper...	D	Salinidad	S	Disco se...
Row0	S/D	Vertical	Polycystina	Spumellaria	N.N.	N.N.	N.N. (radiolario)	40	91.429	R. Bucheli	R. Bucheli	24.6	?	?														
Row1	S/D	Vertical	Globothalam...	Rotalida	Globigerina	N.N.	Globigerina sp.	10	22.857	R. Bucheli	R. Bucheli	24.6	?	?														
Row2	S/D	Vertical	Globothalam...	Rotalida	Globocadri...	N.N.	Globocadrina...	20	45.714	R. Bucheli	R. Bucheli	24.6	?	?														
Row3	S/D	Vertical	Hydrozoa	Siphonophor...	Lensia	N.N.	Lensia sp.	10	22.857	R. Bucheli	R. Bucheli	24.6	?	?														
Row4	S/D	Vertical	Gastropoda	N.N.	N.N.	N.N.	N.N. (larva vel...	20	45.714	R. Bucheli	R. Bucheli	24.6	?	?														
Row5	S/D	Vertical	Polychaeta	Phyllococida	Travisiopsis	N.N.	Travisiopsis sp.	10	22.857	R. Bucheli	R. Bucheli	24.6	?	?														
Row6	S/D	Vertical	Polychaeta	Phyllococida	Tomopteris	N.N.	Tomopteris sp.	10	22.857	R. Bucheli	R. Bucheli	24.6	?	?														
Row7	S/D	Vertical	Ostracoda	Halocyprida	Conchoecia	N.N.	Conchoecia sp.	10	22.857	R. Bucheli	R. Bucheli	24.6	?	?														

base de datos.

Figura 18. Vista preliminar de los datos en el nodo Excel Reader (XLS). Tomado de KNIME Analytics Platform

Estructura de la tabla en herramienta KMINE:

Columns: 22	Column Type	Column Index	Color Handler	Size Handler	Shape Han...
FUENTE	String	0			
ESTACIÓN	String	1			
HORA DE COLECTA	Local Date Time	2			
DÍA	String	3			
MES	String	4			
AÑO	Number (integer)	5			
LATITUD	String	6			
LONGITUD	String	7			
profundidad colecta de muestra	String	8			
TIPO DE ARRASTRE	String	9			
CLASE	String	10			
ORDEN	String	11			
GÉNERO	String	12			
ESPECIE	String	13			
Nombre científico de especie	String	14			
Especie	Number (integer)	15			
Transformados	Number (double)	16			
Responsable del análisis:	String	17			
Llenado por:	String	18			
Temperatura	Number (double)	19			
Salinidad	Number (double)	20			
Disco secchi	String	21			

Figura 19. Estructura de tabla en herramienta KNIME. Tomado de KNIME Analytics Platform

2. Preprocesamiento / limpieza de los datos

Después de que los datos son cargados en la primera etapa de selección de datos usando el nodo Excel Reader en la herramienta, se agregaron dos nodos que ayudarán en el proceso de limpieza y preprocesamiento de los datos.

Como primer paso en esta etapa se eliminaron datos que hacían ruido en el conjunto, para lo cual se usaron los nodos Row Filter y Missing Value.

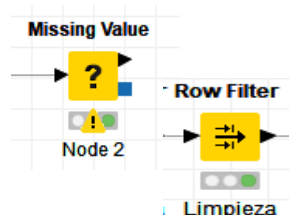


Figura 19. Nodos Missing Value y Row Filter. Tomado de KNIME Analytics Platform

Existen columnas con datos que no serán útiles para la creación del modelo, tales como las columnas relacionadas con los organismos zooplanctónicos (*orden*, *clase*, *especie*, *género*) que contienen como dato “N.N.” que indica “No Nombre”, debido a que estos datos son indispensables, se usó el nodo row filter, el cual eliminó todos los datos “N.N.” Siendo eliminados el 55% de los registros por ruido en su contenido.

Filtered - 2:9 - Row Filter (Limpieza)

File Hilite Navigation View

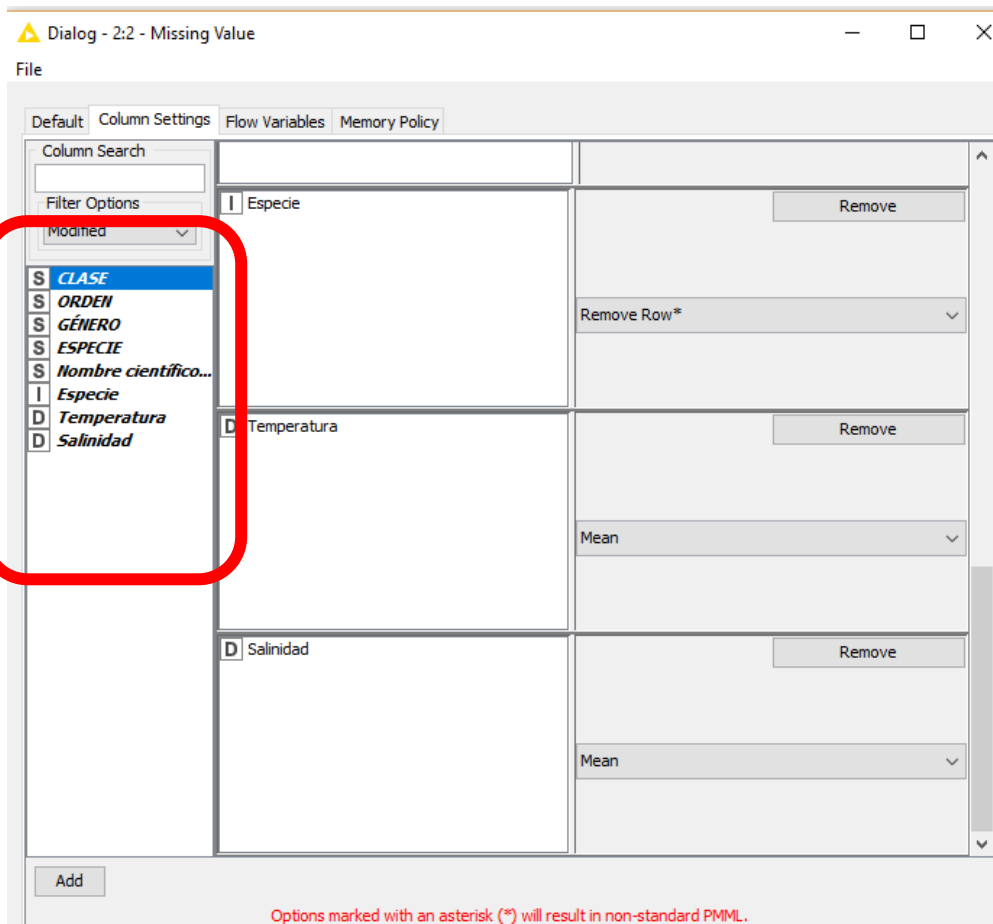
Table "Base de Datos.xlsx [Base de datos]" - Rows: 7375 Spec - Columns: 22 Properties Flow Variables

...	ITUD	LONGITUD	profundi...	TIPO DE...	CLASE	ORDEN	GÉNERO	ESPECIE	Nombre cie...	Especie
...	1' S	81°02.075' W	S/D	Vertical	Hexanauplia	Calanoidea	Calanus	australis	Calanus australis	20
...	1' S	81°02.075' W	S/D	Vertical	Hexanauplia	Calanoidea	Undinula	darwini	Undinula darwini	10
...	1' S	81°02.075' W	S/D	Vertical	Hexanauplia	Calanoidea	Paracalanus	indicus	Paracalanus ind...	60
...	1' S	81°02.075' W	S/D	Vertical	Hexanauplia	Calanoidea	Paracalanus	campaneri	Paracalanus ca...	60
...	1' S	81°02.075' W	S/D	Vertical	Hexanauplia	Calanoidea	Acrocalanus	longicornis	Acrocalanus lon...	20

Figura 20. Tabla de datos después de ejecución de nodo Row Filter. Tomado de KNIME Analytics Platform

Para los registros que no contienen datos se usó el nodo *Missing Value*; las columnas que fueron afectadas son:

Figura 21. Campamentos que serán eliminados



ados en caso de ausencia. Tomado de KNIME Analytics Platform

Como se muestra en la figura son ocho las columnas que serán eliminadas si tiene valores nulos o llenadas con valores determinados dependiendo del análisis de cada uno y de su relación e importancia en este trabajo, lo cual se detallará en la siguiente tabla.

Columna	Acción	Explicación
Clase Orden Género Especie Nombre científico Número de especies	<i>Eliminar fila</i>	Se prefiere eliminar la fila debido a que los principales objetos de estudio son los organismos y si no tenemos información sobre los mismo el registro no aporta
Temperatura Salinidad	<i>Calcular la Media</i>	Debido a la importancia de estos registros, se prefiere no eliminarlos, sino que se autocompleten con la media de los valores existentes.

Tabla 9. Detalle de la limpieza de datos. Autora

Después de la ejecución del tercer nodo y aplicando los cambios perteneciente a la limpieza de los datos se realizó lo que se recomienda en esta etapa que los valores numéricos que se encuentren vacíos sean completados con la media o algún valor estadístico relacionado, siguiendo con la explicación que se detalló en la tabla 9, las columna de temperatura fue afectada 40% debido a su ausencia de valores y la columna de salinidad fue afectada 77% por el mismo motivo, y tal como ya fue mencionado anteriormente estas columnas fueron llenadas con el valor de su media 24.6° y 33.85 UPS respectivamente.

Output table - 2:2 - Missing Value (Preprocesamiento/limpieza)

File Hilite Navigation View

Table "default" - Rows: 7375 | Spec - Columns: 22 | Properties | Flow Variables

SE	ORDEN	GÉNERO	ESPECIE	Nombre ci...	Especie	Transfo...	Respon...	Llenado...	Temper...	Salinida	
...	plia	Calanoida	Calanus	australis	Calanus australis	20	45.714	R. Bucheli	R. Bucheli	24.6	33.85
...	plia	Calanoida	Undinula	darwini	Undinula darwini	10	22.857	R. Bucheli	R. Bucheli	24.6	33.85
...	plia	Calanoida	Paracalanus	indicus	Paracalanus indi...	60	137.143	R. Bucheli	R. Bucheli	24.6	33.85
...	plia	Calanoida	Paracalanus	campaneri	Paracalanus ca...	60	137.143	R. Bucheli	R. Bucheli	24.6	33.85
...	plia	Calanoida	Acrocalanus	longicornis	Acrocalanus lon...	20	45.714	R. Bucheli	R. Bucheli	24.6	33.85
...	plia	Calanoida	Calocalanus	pavoninus	Calocalanus pav...	20	45.714	R. Bucheli	R. Bucheli	24.6	33.85
...	plia	Calanoida	Subeucalanus	pileatus	Subeucalanus ...	60	137.143	R. Bucheli	R. Bucheli	24.6	33.85

Figura 22. Base de datos luego de la etapa de Preprocesamiento y limpieza.
Tomado de KNIME Analytics Platform

3. Transformación / reducción

En la esta etapa, el producto es la vista minable, para lo cual se eliminaron algunas columnas que son redundantes y que se usarán para en el proceso de los datos. Las columnas eliminadas están detalladas en la tabla 10 a continuación:

Columna Eliminada

Detalle

Longitud

Debido a que existen estaciones fijas de trabajo, es redundante con la columna estación.

Latitud

Transformados

Debido a que existe la columna de número de especies, este valor se convierte en redundante y no usado cuando se procesen los datos

Genero	Se eliminan estas columnas debido a que existe la columna de Nombre científico que es la unión de ambas y sería redundante tener las 3 columnas.
Especie	

Fuente

Profundidad

Responsable del análisis

Dato que no será usado en el modelo

Llenado por

Disco sechi

Hora de colecta

Tabla 10. Detalle de eliminación de columnas en etapa de Transformación. Autora.

Para el proceso detallado anteriormente en la herramienta se usó el nodo llamado *Column Filter*, el cual elimina columnas que no serán usadas y al ser ejecutado genera la tabla con las columnas que serán utilizadas en el proceso.

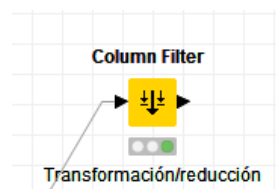


Figura 23. Nodo Column Filter. Tomado de KNIME Analytics Platform

Tal y como ya se indicó anteriormente, el uso de este nodo sirvió para quitar columnas innecesarias en la base de datos que se está usando en el proceso; tal y como se muestra en la figura 22 a continuación.

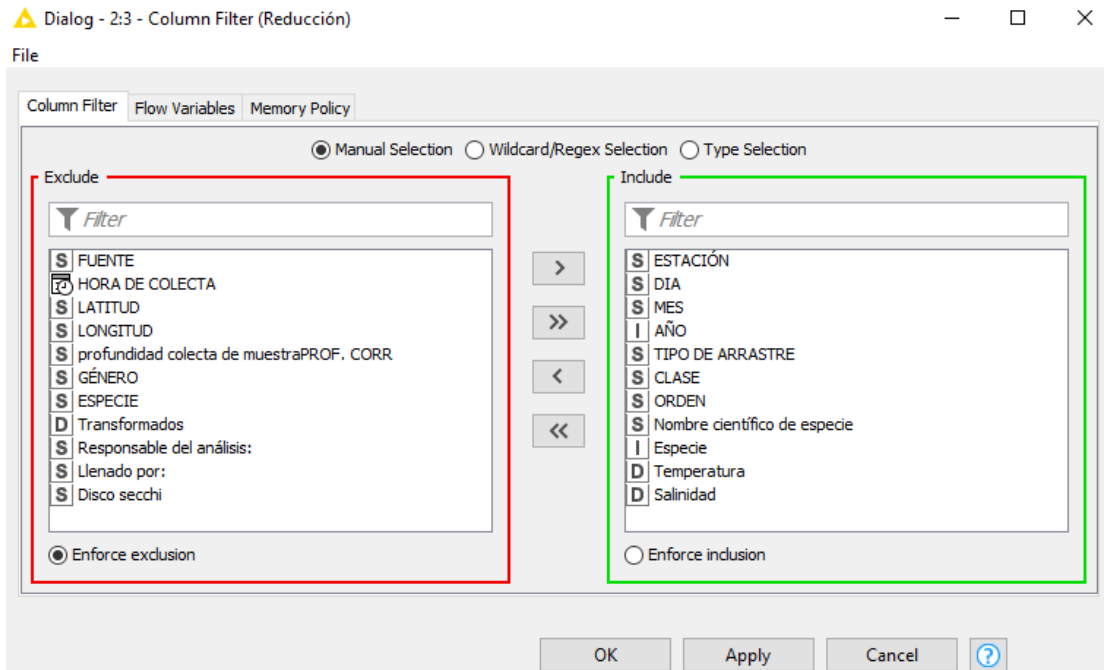


Figura 24. Columnas eliminadas. Tomado de KNIME Analytics Platform

Como se puede ver en la figura 22, las columnas que se encuentran del lado izquierdo y remarcadas en un cuadro rojo son las que tal y como se detalló en la tabla 10, debido a diferentes factores y características fueron eliminadas para que la base de datos que será procesada quede libre de datos con ruidos o redundantes.

ESTACIÓN	DIA	MES	AÑO	TIPO DE...	CLASE	ORDEN	Nombre cien...	Especie	Tempera...	Salinidad
.. Puerto López	19	Abril	2008	Vertical	Hexanaupla	Calanoida	Calanus australis	20	24.6	33.85
.. Puerto López	19	Abril	2008	Vertical	Hexanaupla	Calanoida	Undinula darwini	10	24.6	33.85
.. Puerto López	19	Abril	2008	Vertical	Hexanaupla	Calanoida	Paracalanus indi...	60	24.6	33.85
.. Puerto López	19	Abril	2008	Vertical	Hexanaupla	Calanoida	Paracalanus ca...	60	24.6	33.85
.. Puerto López	19	Abril	2008	Vertical	Hexanaupla	Calanoida	Acrocalanus lon...	20	24.6	33.85
.. Puerto López	19	Abril	2008	Vertical	Hexanaupla	Calanoida	Calocalanus pav...	20	24.6	33.85
.. Puerto López	19	Abril	2008	Vertical	Hexanaupla	Calanoida	Subeucalanus pl...	60	24.6	33.85
.. Puerto López	19	Abril	2008	Vertical	Hexanaupla	Calanoida	Rhincalanus cor...	20	24.6	33.85
.. Puerto López	19	Abril	2008	Vertical	Hexanaupla	Calanoida	Clausocalanus ...	20	24.6	33.85

Figura 25. Tabla generada después de la ejecución del nodo column filter. Tomado de KNIME Analytics Platform

Una vez que se realizó la reducción de columnas que no serán utilizadas en el modelo, se explicará la data que fue discretizada, lo cual está detallado en la siguiente tabla.

Dato (columna de la base de datos)

Detalle

Según lo detallado por los especializados en el tema se tienen los siguientes rangos:

Especie

0 – 100	No abundante
100-1000	Condiciones normales
>1000	Abundancia

Según lo detallado por los especializados en el tema se tienen los siguientes rangos:

Temperatura

TSM < 20°C	Aguas frías
20°C > TSM > 24.5°C	Aguas normales
TSM > 24.5°C	Aguas cálidas

Salinidad

Tipos de masas de aguas superficiales frente a la costa ecuatoriana. Tomado de (Enfield, 1976)

Aguas Subtropicales	35 UPS
Aguas Tropicales	<33 UPS
Aguas Ecuatoriales (condiciones normales)	33 UPS a 34,9 UPS
Aguas Costeras de Humboldt	>34.9 UPS

Tipo especie

De las especies encontradas mensualmente en cada estación se realizó la siguiente clasificación según las características del agua presentadas:

especie Indicador niño	<ul style="list-style-type: none"> • Temperatura=Aguas cálidas ^ Especie=
------------------------	--

	Abundante	
	<ul style="list-style-type: none"> • Temperatura=Aguas Frías \wedge Especie=No abundante 	
especie Indicador niña	<ul style="list-style-type: none"> • Temperatura=Aguas cálidas \wedge Especie=No abundante • Temperatura=Aguas Frías \wedge Especie Abundante 	
Condiciones normales	Especie= Normales	Condiciones

Estado fenómeno Según las características del mar y la distribución y abundancia de los organismos se realiza la siguiente clasificación:

Fenómeno_Niño	<ul style="list-style-type: none"> • Tipo de Especie= Especie_indicador_niño \wedge Especie =Abundante \wedge Temperatura=Aguas cálidas • Tipo de Especie=Especie_indicador_niño o \wedge Especie=Abundante \wedge Temperatura=Aguas Frías • Tipo de Especie= Especie_indicador_niña \wedge Especie\$ =No abundante \wedge Temperatura=Aguas cálidas
Fenómeno_ Niña	<ul style="list-style-type: none"> • Tipo de Especie= =Especie_indicador_niño \wedge Especie= No abundante \wedge Temperatura= "Aguas Frías

		<ul style="list-style-type: none"> Tipo de Especie= Especie_indicador_niña Especie= Abundante Temperatura=Aguas Frías
	Condiciones_Normales	<ul style="list-style-type: none"> Tipo de Especie= Especie_Condiciones_Normales Especie= Condiciones Normales
Tipo de Arrastre	Los valores que tomara son: V: Vertical S: Superficial	

Tabla 11. Detalle de datos a discretizar. Autora

Para realizar lo detallado en la tabla 11, se utilizaron dos nodos que brinda la herramienta en uso, los cuales permiten realizar la transformación de los datos a nominales, categóricos. Los nodos usados son:

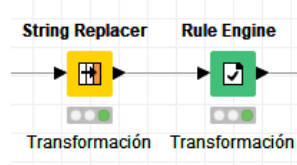


Figura 26. Nodos usados para transformación de datos. Tomado de KNIME Analytics Platform

Para la transformación a un valor nominal de la columna *Tipo de arrastre*, se utilizó el nodo *String Replacer*, el cual a tal y como se indicó en la tabla 11, sus valores fueron abreviados, y después de su transformación quedó de la siguiente manera:

Input with replaced values - 2:12 - String Replacer (Transformación)

File Hilite Navigation View

Table "default" - Rows: 7375 Spec - Columns: 11 Properties Flow Variables

...	S ESTACIÓN	S DIA	S MES	I AÑO	S TIPO DE...	S CLASE	S ORDEN	S Nombre cie...	I Especie	D Tei
...	Esmeraldas	20	Abril	2008	V	Hexanauplia	Harpacticoida	Euterpina acutif...	20	27.9
...	Esmeraldas	20	Abril	2008	V	Hexanauplia	Harpacticoida	Clytemnestra sc...	10	27.9
...	Esmeraldas	20	Abril	2008	S	Branchiopoda	Diplostraca	Evadne tergestina	60	27.9
...	Esmeraldas	20	Abril	2008	S	Hexanauplia	Calanoida	Canthocalanus ...	40	27.9

Figura 27. Datos transformados por nodo String Replacer. Tomado de KNIME Analytics Platform

El nodo *Rule Engine* se usó para la transformación de las columnas *Temperatura* y *Especie* dándole valores categóricos a sus datos, dependiendo del rango ya detallado para cada una en la tabla 11.

Después de que los datos fueron discretizados y reducidos se puede mostrar la vista minable obtenida a través del proceso de transformación que se ha llevado a cabo.

Classified values - 2:14 - Rule Engine (Transformación)

File Hilite Navigation View

Table "default" - Rows: 7375 Spec - Columns: 11 Properties Flow Variables

...	S ESTACIÓN	S DIA	S MES	I AÑO	S TIPO DE...	S CLASE	S ORDEN	S Nombre cien...	S Especie	S Temperat...	D Salinidad
..	Esmeraldas	23	Mayo	2008	V	Hexanauplia	Cyclopoida	Oncaea venusta	No abundante	Aguas cálidas	33.85
..	Esmeraldas	23	Mayo	2008	V	Hexanauplia	Cyclopoida	Ditrichocorycae...	No abundante	Aguas cálidas	33.85
..	Esmeraldas	23	Mayo	2008	V	Hexanauplia	Cyclopoida	Farranulla rostr...	No abundante	Aguas cálidas	33.85
..	Esmeraldas	23	Mayo	2008	V	Hexanauplia	Cyclopoida	Farranulla gracilis	No abundante	Aguas cálidas	33.85
..	Esmeraldas	23	Mayo	2008	V	Hexanauplia	Harpacticoida	Microsetella rosea	No abundante	Aguas cálidas	33.85
..	Esmeraldas	23	Mayo	2008	S	Globothala...	Rotaliida	Globigerinoides...	No abundante	Aguas cálidas	33.85
..	Esmeraldas	23	Mayo	2008	S	Gastropoda	Pteropoda	Limacina trochif...	Abundante	Aguas cálidas	33.85
..	Esmeraldas	23	Mayo	2008	S	Branchiopoda	Diplostraca	Evadne tergestina	No abundante	Aguas cálidas	33.85
..	Esmeraldas	23	Mayo	2008	S	Ostracoda	Halocyprida	Conchoecia obt...	No abundante	Aguas cálidas	33.85
..	Esmeraldas	23	Mayo	2008	S	Hexanauplia	Calanoida	Canthocalanus ...	Condiciones Normales	Aguas cálidas	33.85
..	Esmeraldas	23	Mayo	2008	S	Hexanauplia	Calanoida	Undinula vulgaris	Condiciones Normales	Aguas cálidas	33.85
..	Esmeraldas	23	Mayo	2008	S	Hexanauplia	Calanoida	Paracalanus ind...	Condiciones Normales	Aguas cálidas	33.85

Figura 27. Vista Minable. Tomado de KNIME Analytics Platform

La vista minable generada por la herramienta, se la obtuvo después del proceso de transformación y reducción necesario en la metodología en uso, para lo cual se realizó un diccionario de datos, detallando cada columna de la base de datos utilizada, lo cual se muestra en la tabla 12 a continuación.

Columnas de la base de datos	Tipo de dato	Descripción	Valores posibles
Estación	Nominal	Lugar del muestreo	Puerto López Salinas Esmeraldas

				Santa Clara Punta Galera
Tipo de arrastre	Nominal	Método de colecta de muestra		Vertical (V), Superficial (S)
Clase	Nominal	Grupos de organismos		Ver tabla 1 en Anexo 1
Orden	Nominal	Familia a la que pertenece el organismo		Ver tabla 1 en Anexo 1
Tipo de especie	Nominal	Clasificación de organismos según características del agua		Especie_Indicador_niño Especie_Indicador_niña Especie_Condiciones_normales
Especie	Nominal	Número de organismos encontrados		No abundante Condiciones normales Abundante
Temperatura	Nominal	Valor de la temperatura superficial del mar		Aguas frías Aguas normales Aguas cálidas
Salinidad	Nominal	Valor de la salinidad presente en el mar		Aguas Subtropicales Aguas Tropicales Aguas Ecuatoriales(condiciones normales) Aguas Costeras de Humboldt
Estado fenómeno	Nominal	Clasificación de tipo de fenómeno según características presentadas.		Fenómeno_Niño Fenómeno_Niña Condiciones_Normales

Tabla 12. Diccionario de datos Vista minable. Autora

4. Minería de Datos

El modelo que se generará es de tipo *predictivo*, y las técnicas que serán usadas son de *clasificación*, por lo cual la variable de clasificación que se usará para el mismo es *Estado fenómeno*, la cual indicara según reglas generadas, la presencia o ausencia de algún evento ENOS. Para lo cual los algoritmos que se usaran son:

Arboles de Clasificación: Un árbol consiste esencialmente en un conjunto secuencial de condiciones y acciones que relacionan unos determinados factores con un resultado o decisión. Es un método de clasificación supervisada, lo que quiere decir que cuenta con datos ya clasificados a priori de los que extraerá el conocimiento. En primer lugar el

algoritmo analiza los datos proporcionados en busca de patrones y usa los resultados de este análisis para definir la secuencia y condiciones para la creación del modelo de clasificación. (Porras, 2014)

El algoritmo que se empleó para el árbol de decisión obtenido fue C4.5, el cual en la herramienta en uso se lo conoce con el nombre del nodo J48 que es parte de los nodos pertenecientes a la extensión Weka que la herramienta en uso (KNIME) posee. Para generar el modelo predictivo a partir de este algoritmo también se usó el nodo Weka Predictor el cual nos muestra el árbol de clasificación generado.

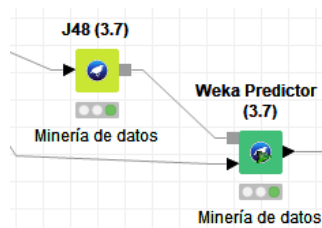
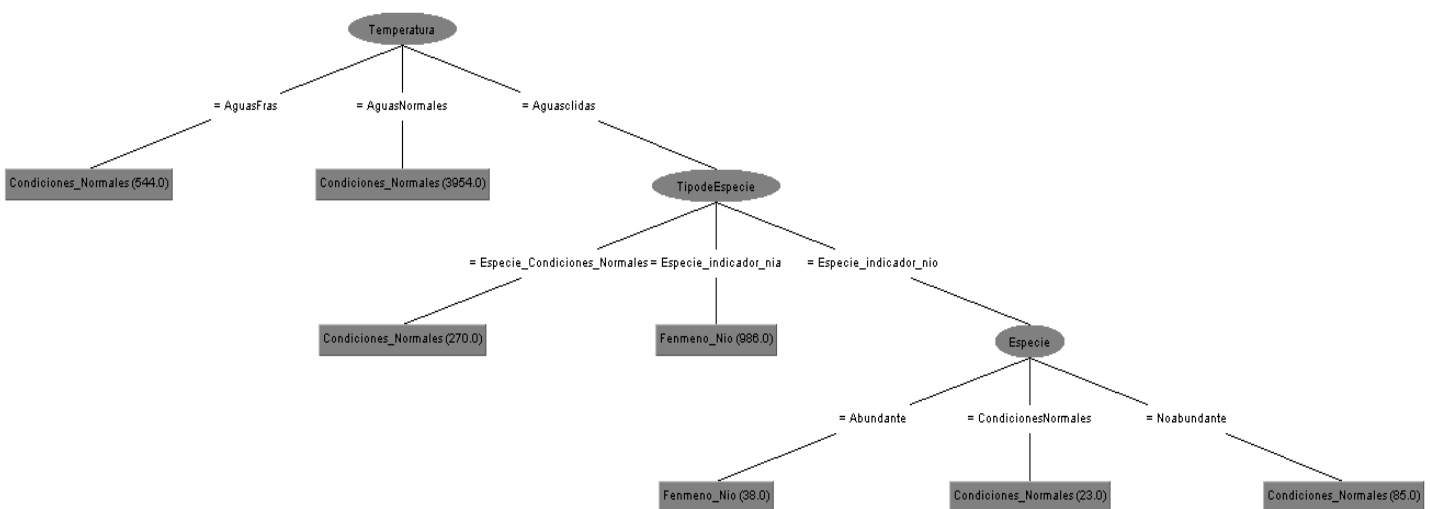


Figura 28. Nodos Para implementar algoritmo C4.5. Tomado de KNIME Analytics Platform

Resultados de Algoritmo C 4.5 (Nodo J48)

Mediante este algoritmo se obtuvo como resultado un árbol de clasificación basado en las reglas que proporcionan las variables que son



parte de la vista minable.

Figura 29. Arbol de clasificación obtenido con algoritmo J48. Tomado de KNIME Analytics Platform

Detalle árbol:

La clasificación de reglas que generó el algoritmo para crear el árbol de decisiones parte de la temperatura del agua, si sus condiciones son cálidas, se procede a revisar el tipo de especies presentes, teniendo las especies condiciones normales, que como su nombre lo denota indica que no hay presencia de evento ENOS, se mantienen las condiciones normales, para la especie indicador de niña en poco abundancia y casi ausente en aguas cálidas indica presencia del Fenómeno_Niño, y la especie indicador de niño cuando es abundante indica presencia de Fenómeno_Niño, si existe la especie indicador niña de manera no abundante esto también indica presencia de Fenómeno_niño.

Reglas de clasificación: Las reglas permiten expresar disyunciones de manera más fácil que los árboles y tienden a preferirse con respecto a los árboles por tender a representar "pedazos" de conocimiento relativamente independientes. (Robles & Sotolongo, 2013). Las técnicas de Inducción de Reglas permiten generar y contrastar árboles de decisión, o reglas y patrones a partir de los datos de entrada. La información de entrada será un conjunto de casos en que se ha asociado una clasificación o evaluación a un conjunto de variables o atributo (Ruiz, 2009).

El algoritmo con el que se implementaron las reglas de clasificación es el PRISM, el cual también forma parte de la extensión Weka en la herramienta en uso, y su nombre es el mismo que el algoritmo implementado, acompañado del nodo Weka Predictor se generan las reglas de clasificación para el modelo predictivo.

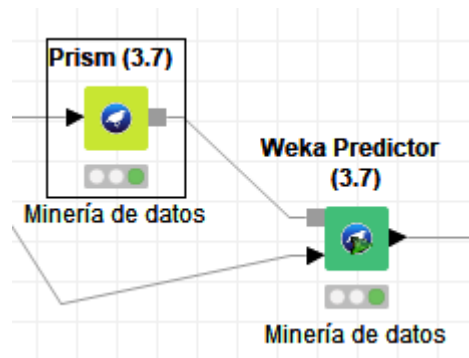


Figura 30. Nodos Para implementar algoritmo PRISM. Tomado de KNIME Analytics Platform

Detalle de Reglas:

Las reglas de clasificación obtenidas son las siguientes:

```

Weka Node View - 0:20 - Prism (3.7) (Minería de datos)
File
Weka Output
Prism rules
-----
If Temperatura = AguasNormales then Condiciones_Normales
If Temperatura = AguasFras then Condiciones_Normales
If Especie = CondicionesNormales then Condiciones_Normales
If ORDEN = val2 then Condiciones_Normales
If Salinidad = AguasTropicales
  and Especie = Noabundante then Condiciones_Normales
If ORDEN = Anthoathecata then Fenmeno_Nio
If Temperatura = Aguasclidas
  and TipodeEspecie = Especie_indicador_nia then Fenmeno_Nio
If Especie = Abundante
  and Temperatura = Aguasclidas then Fenmeno_Nio

```

Figura 29. Reglas de clasificación obtenido con algoritmo PRISM. Tomado de KNIME Analytics Platform

El modelo generado, se visualiza de la siguiente manera en la herramienta:

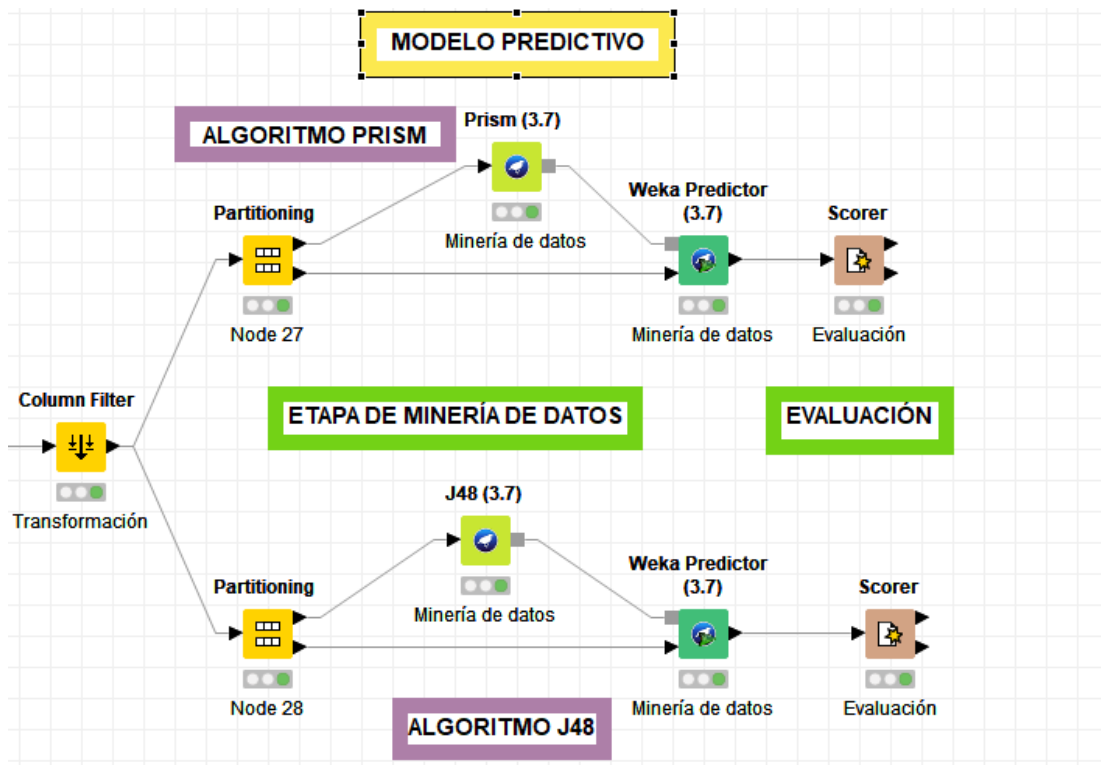


Figura 30. Modelo predictivo. Tomado de KNIME Analytics Platform

5. Interpretación y Evaluación

La evaluación del modelo generado se realizó por medio de la herramienta con el nodo Score, el cual mediante la matriz de confusión muestra los resultados de evaluación del modelo.

Matriz de confusión (Algoritmo C4.5 y PRISM)

Algoritmo	Matriz de Confusión
------------------	----------------------------

C4.5 (Árbol de clasificación)

Estado Fe...	Fenómeno_Niño	Condiciones_Normales
Fenómeno_...	1296	0
Condiciones...	107	5892

Correct classified: 7.188 Wrong classified: 107
Accuracy: 98,533 % Error: 1,467 %
Cohen's kappa (κ) 0,951

PRISM (Reglas de clasificación)

Estado Fenomeno \ P...	Fenómeno_Niño	Condiciones_Normales
Fenómeno_Niño	1242	49
Condiciones_Normales	1	6003

Correct classified: 7.245 Wrong classified: 50
Accuracy: 99,315 % Error: 0,685 %
Cohen's kappa (κ) 0,976

Con error en rango aceptable, por debajo del 5%, se generaron los resultados por medio de los algoritmos usados ya antes mencionados.

Para la evaluación del modelo se tienen las siguientes variables:

Precisión y exactitud: Aproximación que tiene el modelo a los valores reales y poca dispersión de los mismos.

Sensibilidad y especificidad: La capacidad que tiene el modelo de diferenciar los casos que son negativos de los casos que son positivos.

Para el caso del modelo se tiene los siguientes valores:

	D Recall	D Precision	D Sensitivity	D Specifity	D F-meas...	D Accuracy	D Co
1		0.924	1	0.982	0.96	?	?
0.982	1		0.982	1	0.991	?	?
?	?	?	?	?	?	0.985	0.951

tives	D Recall	D Precision	D Sensitivity	D Specifity	D F-meas...	D Accuracy	D Co
	0.962	0.999	0.962	1	0.98	?	?
1		0.992	1	0.962	0.996	?	?
?	?	?	?	?	?	0.997	0.971

Figura 31. Variables de evaluación en Matriz de confusión. Tomado de KNIME Analytics Platform

Tal y como se muestra en la figura 30, los valores en las variables de evaluación respaldan de manera positiva el modelo, indicando con dichos valores que es preciso, exacto, específico y sensible (Alta tasa de predicciones positivas).

Interpretación de reglas y Modelo Predictivo:

Se toman como variables de las reglas los siguientes campos:

Temperatura, especie, tipo de especie y orden.

Después de los algoritmos aplicados, estas son las variables que determinan el modelo, debido a que son factores determinantes y valores relevantes que las mismas brindan para poder indicar la ocurrencia del ciclo ENOS.

Las reglas generadas a partir de los datos procesados y que forman parte del modelo propuesto tienen como resultado 2 condiciones en el estado del fenómeno que es la variable de clasificación; los cuales son las siguientes:

Fenómeno del niño

Aguas cálidas + Especie Abundante + Especie indicador niño
Aguas cálidas + Especie No abundante + Especie indicador niña
Anthoathecata + Especie Abundante

Condiciones Normales

Temperatura = AguasNormales
Especie = CondicionesNormales

Este conjunto de reglas son el resultado del uso de las técnicas de minería de datos, luego de un proceso completo, basado en la metodología KDD las reglas expuestas generan el modelo predictivo de la ocurrencia de eventos extremos de la variabilidad climática en la zona costera del Ecuador.

CONCLUSIONES

En el presente trabajo se pudo realizar un breve estudio y reconocimiento del ecosistema marino- costero el cual está directamente ligado con los organismos zooplanctónicos que son parte fundamental del mismo.

Una vez que se tuvo el conocimiento ya antes mencionado, se pudo ejecutar paso a paso las etapas de la metodología KDD, la cual, mediante sus procedimientos en cada paso, se logró procesar los datos hasta obtener una base libre de registros vacíos, nulos o que sean ruido. La etapa de la transformación de los datos se la realizó por medio de la discretización de los datos, y también la eliminación de columnas con datos considerados redundantes o poco relevantes para el modelo, luego de estas dos etapas de la metodología se obtuvo la vista minable con los siguientes campos: Estación, Tipo de arrastre Clase, Orden, Tipo de especie, Especie, Temperatura, Salinidad, Estado fenómeno.

Los algoritmos usados para el modelo son de clasificación debido a que el modelo que se generó es predictivo y estos algoritmos son propios de este tipo de modelos, debido a que ya se tiene una variable de clasificación y se encarga de agrupar los datos, se eligieron 2 algoritmos para el caso de este trabajo, el algoritmo C4.5 que es un árbol de clasificación, y el algoritmo PRISM que es de reglas de clasificación a través de los cuales se obtuvieron reglas y patrones con lo que se generó el modelo, a partir de la variable de clasificación definida para la predicción que fue el campo Estado del Fenómeno.

A partir de los resultados de los algoritmos usados, se generó el modelo predictivo propuesto el cual para ambos la efectividad del mismo fue superior al 95% tal y como se mostró en la etapa de evaluación del modelo, por lo que con estos valores se puede decir que el modelo cumple con el nivel de efectividad, exactitud, sensibilidad. El modelo generado da como resultado dos estados del ciclo ENOS: Fenómeno del niño y condiciones

normales, a partir de las reglas generadas, se puede decir que la distribución de los organismos son parte de la predicción de estos estados.

El ciclo ENOS es un tema de mucho interés para autoridades, por lo cual el modelo generado, es útil para toma de decisiones con respecto planificación de trabajos, investigaciones y/o planes de contingencias que a partir del mismo se puedan realizar.

RECOMENDACIONES

En el análisis se ha evidenciado que los datos procedentes de los estudios e investigaciones realizadas, provienen de la zona norte y centro de la zona costera en su mayoría, por lo que para que los resultados de las predicciones tengan las tres condiciones del Ciclo ENOS, Fenómeno del niño, fenómeno de la niña y condiciones normales. Para lo cual es necesario que la base de datos siga siendo completada con datos de todas las zonas de la región costera de manera progresiva.

Se sugiere que en las reuniones mensuales realizadas en el Comité ERFEN sea implementado el modelo como parte de las herramientas usadas para la toma de decisiones con respecto al tema de ocurrencia del ciclo ENOS.

Los datos y las reglas generadas, son útiles para estudios relacionados con la productividad de la pesca, y la conservación de especies, ya que dichos trabajos involucran a los organismos zooplanctónicos, ya que como son bioindicadores también son parte de la toma de decisiones para el ámbito pesquero.

BIBLIOGRAFÍA

- Agencia Estatal de Metereología del Gobierno de España. (2016). *MeteoGlosario Visual*. Obtenido de Diccionario ilustrado de Metereología:
https://meteoglosario.aemet.es/es/termino/681_variabilidad-climatica
- Alberto, J. A. (2009). *Unne*. Recuperado el 03 de 06 de 2019, de Unne:
http://hum.unne.edu.ar/academica/departamentos/dptogeog/catedras/biogeografia/fac_temperatura.pdf
- Bautista-Cabezas, S. V. (2015). *RepositorioUG*. Recuperado el 06 de 06 de 2019, de RepositorioUG:
<http://repositorio.ug.edu.ec/handle/redug/11526>
- Bergman, J. (30 de agosto de 2001). *Ventanas al Universo*. (J. Bergman, Editor) Recuperado el 31 de 05 de 2019, de Ventanas al Universo:
https://www.windows2universe.org/earth/Water/dissolved_salts.html&lang=sp
- Bergman, J. (16 de 02 de 2011). *Ventanas Al Universo*. (J. Bergman, Editor) Recuperado el 31 de 05 de 2019, de VENTANAS AL UNIVERSO:
<https://www.windows2universe.org/earth/Water/temp.html&lang=sp>
- Boltovskoy, E., & Theyer, F. (1970). Foraminiferos recientes de Chile central. *Revista del Museo Argentino de Ciencias Naturales "Bernardino Rivadavia" e Instituto Nacional de Investigación de las Ciencias Naturales., Hidrobiología(2)*, 279-379. Recuperado el 22 de 05 de 2019
- Caridad-Ocerin, J. M. (Julio-Diciembre de 2011). La Minería De Datos: Análisis De Bases De Datos En La Empresa. *Boletín de la Real Academia de Córdoba de Ciencias, Bellas Letras y Nobles Artes(141)*, 357-369. Recuperado el 06 de 06 de 2019, de
<http://hdl.handle.net/10396/6657>

- Cávez-López, R., Rocha-Ramírez, A., & Ramírez-Rojas, A. (10 de 08 de 2005). Cambios en los ensamblajes de peces del istema launar de Alvarado (sla), Veracruz, México. *Revista Digital Universitaria*, 6(8). Recuperado el 03 de 06 de 2019, de <http://www.revista.unam.mx/vol.6/num8/art79/int79.htm>
- Centro Regional de Información sobre Desastres (CRID). (Abril de 2000). *Centro Regional de Información sobre Desastres (CRID)*. (C. R. (CRID), Ed.) Recuperado el 03 de 06 de 2019, de Centro Regional de Información sobre Desastres (CRID): <https://relaciger.net/revista/documentos/VCD.pdf>
- Cifuentes, J. L., Torres-García, M. d., & Mondragón, M. F. (1997). I. LOS Ciclos Orgánicos Y Niveles Tróficos En Los Océanos. En J. L. Cifuentes, M. d. Torres-García, & M. F. Mondragón, *El Océano Y Sus Recursos Vii. Flujos De Energía En El Mar: Reproducción Y Migraciones* (Segunda ed.). México DF, México DF, México: Ciencia para Todos. Recuperado el 30 de 05 de 2019, de <http://bibliotecadigital.ilce.edu.mx/sites/ciencia/volumen2/ciencia3/063/htm/toc.html>
- Cifuentes-Lemus, J. L., Torres-García, P., & Frías-M, M. (1995). La Temperatura En El Mar. En J. L. Cifuentes-Lemus, P. Torres-García, M. Frías-M, J. L. Cifuentes-Lemus, P. Torres-García, & M. Frías-M (Edits.), *El Océano Y Sus Recursos Iii. Las Ciencias Del Mar: Oceanografía Física, Matemáticas E Ingeniería* (Primera ed.). México, México, México: Ciencia para Todos. Recuperado el 31 de 05 de 2019, de http://bibliotecadigital.ilce.edu.mx/sites/ciencia/volumen1/ciencia2/17/htm/sec_6.html
- Comunidad Andina. (2009). El Niño Oscilación del Sur -ENOS. *Por encima de la Tierra*, 150-155. Recuperado el 07 de junio de 2019, de http://www.comunidadandina.org/predecan/atlasweb/chapters/inicio/pdf/13_El_Nino_y_La_Nina.pdf

- Cortés, A. F. (febrero de 2000). *SIDAR*. Recuperado el 04 de julio de 2019, de SIDAR: <http://www.sidar.org/recur/desdi/traduc/es/visitable/nuevos/Rapido.htm>
- Coto, C. F. (26 de 08 de 2015). *DGCS - Dirección General de Comunicación Social*. Recuperado el 29 de 05 de 2019, de DGCS - Dirección General de Comunicación Social: http://www.dgcs.unam.mx/boletin/bdboletin/2015_491.html
- Cruz, M. (2012). Preferencia Y Rangos De Tolerancia A La Temperatura Y Salinidad De Los Pterópodos Y Heterópodos Frente A La Costa Ecuatoriana. *Acta Oceanográfica Del Pacífico*, 17(1). Recuperado el 28 de 05 de 2019, de https://www.inocar.mil.ec/web/phocadownloadpap/actas_oceanograficas/acta17/OCE1701_6.pdf
- Curso Geografía del Mar. (s.f.). *Curso Geografía del Mar*. Recuperado el 31 de 05 de 2019, de Curso Geografía del Mar: http://www7.uc.cl/sw_educ/geo_mar/html/h321.html
- Curso Geografía Del Mar. (s.f.). *Curso Geografía Del Mar*. Recuperado el 31 de 05 de 2019, de Curso Geografía Del Mar: http://www7.uc.cl/sw_educ/geo_mar/html/h332.html
- Dalen, D. B., & Meyer., W. J. (12 de septiembre de 2006). *Manual de técnica de la investigación educacional*. Recuperado el 07 de junio de 2019, de Manual de técnica de la investigación educacional: <https://noemagico.blogia.com/2006/091301-la-investigaci-n-descriptiva.php>
- Díaz-Bravo, L., Torruco-García, U., Martínez-Hernández, M., & Varela-Ruiz, M. (septiembre de 2013). La entrevista, recurso flexible y dinámico. (L. Díaz-Bravo, U. Torruco-García, M. Martínez-Hernández, & M. Varela-Ruiz, Edits.) *Investigación en educación médica*, 2(7), 162-167. Recuperado el 07 de junio de 2019, de http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S2007-50572013000300009

- El Colegio De La Frontera Sur. (11 de 08 de 2017). *ECOSUR*. Recuperado el 29 de 05 de 2019, de ECOSUR: <https://www.ecosur.mx/zooplancton-pteropodos-y-su-importancia-en-un-oceano-cambiante/>
- Enfield, D. (1976). *Manual de prácticas de Oceanografía Física*. Recuperado el 03 de 06 de 2019, de <http://repositorio.ug.edu.ec/bitstream/redug/807/3/Pet%C3%B3podos%20y%20Heter%C3%B3podos%20como%20biondicadores%20del%20evento%20El%20Ni%C3%B1o%20Parte%202.pdf>
- Espinoza-A, J. (1996). El Niño y sus implicaciones sobre el medio ambiente. (J. Espinoza-A, Ed.) *Acta Oceanográfica del Pacífico*, 8(1), 115-134. Recuperado el 04 de 06 de 2019, de <https://www.oceandocs.org/bitstream/handle/1834/2224/El%20Ni%C3%B1o%20y%20sus%20implicaciones%20sobre%20el%20medio%20ambiente.pdf?sequence=1>
- Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). Knowledge Discovery and Data Mining: *KDD-96 Proceedings*. Recuperado el 23 de 05 de 2019, de <https://www.aaai.org/Papers/KDD/1996/KDD96-014.pdf>
- Flores-Lagla, G. A., Cadena-Moreano, J. A., Quinatoa-Arequipa, E. E., & Quishpe, M. W. (31 de 01 de 2019). Minería de datos como herramienta estratégica. (G. A. Flores-Lagla, J. A. Cadena-Moreano, E. E. Quinatoa-Arequipa, & M. W. Quishpe, Edits.) *Recimundo (Revista científica Mundo de la Investigación y el Conocimiento)*, 3(1), 955-970. doi:[https://doi.org/10.26820/recimundo/3.\(1\).enero.2019.955-970](https://doi.org/10.26820/recimundo/3.(1).enero.2019.955-970)
- García, B., & I., R. (Mayo de 2011). Reseña del Modelo de Prototipo y Herramientas Case. *Reseña del Modelo de Prototipo y Herramientas Case*. Recuperado el 17 de julio de 2019, de <http://gestionrrhusm.blogspot.com/2011/05/modelo-de-prototipo.html>
- García-Herrero, J., & Molina-López, J. M. (2012). Técnicas De Análisis De Datos. *Técnicas De Análisis De Datos*. Madrid, España, España.

Recuperado el 11 de junio de 2019, de <http://ocw.uc3m.es/ingenieria-informatica/analisis-de-datos/libroDataMiningv5.pdf>

Gutiérrez-Pacherres, J. J. (2017). Propuesta De Una Metodología De Extracción De Conocimientos A Partir De Datos De Las Prestaciones Del Seguro Integral De Salud En La Región Piura En El Año 2016. *Propuesta De Una Metodología De Extracción De Conocimientos A Partir De Datos De Las Prestaciones Del Seguro Integral De Salud En La Región Piura En El Año 2016*. Piura, Piura, Perú. Recuperado El 11 De Junio De 2019, De Http://Repositorio.Uladech.Edu.Pe/Bitstream/Handle/123456789/2868/Mineria_De_Datos_Servicio_De_Salud_Gutierrez_Pacherres_Juan_Jesus.pdf?sequence=1&isAllowed=y

Heink, U., & Kowarik, I. (Mayo de 2010). What are indicators? On the definition of indicators in ecology and environmental planning. *Ecological Indicators*, 10(3), 584-593. Recuperado el 29 de 05 de 2019, de <https://www.sciencedirect.com/science/article/pii/S1470160X09001575>

Henriquez-Daza, M. (2001). ¿Un nuevo fenómeno del Niño? (M. Henriquez-Daza, Ed.) *Revista Entornos*, 2(14), 9-18. Recuperado el 04 de 06 de 2019, de <https://dialnet.unirioja.es/servlet/articulo?codigo=6936969>

Hernández-Sampieri, D. R., Fernández-Collado, D. C., & Baptista-Lucio, D. P. (2005). *Metodología de la investigación* (Cuarta ed.). (M. G. Hill, Ed.) México, México D.F, México: Mc. Graw Hill. Recuperado el 07 de Junio de 2019, de https://s3.amazonaws.com/academia.edu.documents/38758233/sampieri-et-al-metodologia-de-la-investigacion-4ta-edicion-sampieri-2006_ocr.pdf?response-content-disposition=inline%3B%20filename%3DSampieri-et-al-metodologia-de-la-investi.pdf&X-Amz-Algorithm=AWS4

Hernans-Killam, L., & Daou, D. (08 de septiembre de 2001). *Spitzer*. (L. Hernans-Killam, D. Daou, Editores, Spitzer, Productor, & Spitzer)

Recuperado el 31 de 05 de 2019, de Spitzer:
http://legacy.spitzer.caltech.edu/espanol/edu/thermal/temperature_sp_06sep01.html

INOCAR. (s.f.). *INOCAR*. Recuperado el 22 de 05 de 2019, de INOCAR:
https://www.inocar.mil.ec/modelamiento/elnino/nino_generalidades.php

Instituto Nacional de Pesca. (s.f.). *Instituto Nacional de Pesca*. Recuperado el 30 de 05 de 2019, de Instituto Nacional de Pesca:
<https://www.institutopesca.gob.ec/quienes-somos/>

Junta de Andalucía. (2009). Junta de Andalucía. *REDIAM*. Recuperado el 03 de 06 de 2019, de Junta de Andalucía:
http://www.juntadeandalucia.es/medioambiente/portal_web/rediam/contenidos_ordenacion/PDF/FICHA_SST.pdf

Krall, C. (2006). *Minería de datos (data mining). Qué es y para qué sirve. (2ª parte) (DV00106A)*. Recuperado el 06 de 06 de 2019, de Aprendeaprogramar.com:
<https://www.aprenderaprogramar.com/attachments/article/258/DV00106A%20Mineria%20de%20datos%20data%20mining%20modelos%20tecnicas%20herramientas%202.pdf>

Landa, J. (19 de febrero de 2016). *Tratamiento de los datos*. (J. Landa, Editor, J. Landa, Productor, & Javier Landa) Recuperado el 07 de junio de 2019, de Fcojlanda.me: <http://fcojlanda.me/es/ciencia-de-los-datos/kdd-y-mineria-de-datos-espanol/>

Lara Alfredo, M. M. (2014). *Herramientas De Minería De Datos. Herramientas De Minería De Datos*. Huejutla, México, México. Recuperado el 17 de Julio de 2019, de https://www.uaeh.edu.mx/investigacion/productos/6536/herramientas_de_mineria_de_datos.pdf

Lemus, J. L., García, M. D., & M, M. F. (1997). *Propiedades Químicas Del Agua De Mar: Salinidad, Clorinidad Y Ph*. En J. L. Lemus, M. D.

García, M. F. M, J. L. Lemus, M. D. García, & M. F. M (Edits.), *El Océano Y Sus Recursos li. Las Ciencias Del Mar: Oceanografía Geológica Y Oceanografía Química* (Segunda Ed.). México, México, México: Ciencia Para Todos. Recuperado el 31 de 05 de 2019, de http://bibliotecadigital.ilce.edu.mx/sites/ciencia/volumen1/ciencia2/12/html/sec_2.html

LOGICALIS. (12 de 01 de 2015). *LOGICALIS*. Recuperado el 06 de 06 de 2019, de LOGICALIS: <https://blog.es.logicalis.com/analytics/modelos-de-data-mining-y-las-herramientas-mas-usadas>

López-Cano, J. L. (1986). *Método e hipótesis científicos* (3 ed.). México, México, México: Trillas. Recuperado el 07 de Junio de 2019, de https://books.google.com.ec/books/about/M%C3%A9todo_e_hip%C3%B3tesis_cient%C3%ADficos.html?id=OSwKOgAACAAJ&redir_esc=y

Luna, G. L. (Septiembre de 2011). Minería de datos: cómo hallar una aguja en un pajar. *Ciencia*, 18-28. Recuperado el 28 de 05 de 2019, de https://www.amc.edu.mx/revistaciencia/images/revista/62_3/PDF/mineria_aguja.pdf

Luzuriaga, M. (2014). *INOCAR*. Recuperado el 22 de 05 de 2019, de INOCAR: https://www.inocar.mil.ec/web/phocadownloadpap/actas_oceanograficas/acta19_1/OCE1901_7.pdf

Luzuriaga-Villareal, M. (2015). Foraminíferos Planctónicos Vivos En Condiciones Normales Del Mar. *Acta Oceanografica Del Pacifico*, 20(1). Recuperado el 28 de 05 de 2019, de <https://www.inocar.mil.ec/web/index.php/publicaciones/actas-oceanograficas/category/29-acta-oceanografica-del-pacifico-vol-20-n-1-2015>

Manzano, G., Montesano, R., & Zúñiga, L. (Agosto de 2017). Análisis, diseño e implementación de algoritmos. *Apunte electrónico*, 1-121. Ciudad de México, México DF, México: Universidad Nacional Autónoma de

México. Recuperado el 12 de junio de 2019, de http://fcasua.contad.unam.mx/apuntes/interiores/docs/20181/informatica/1/LI_1164_06097_A_Analisis_Disenio_Implantacion_Algoritmos_Plan2016.pdf

Marínez, I. (13 de 01 de 2016). *UPM*. Recuperado el 31 de 05 de 2019, de UPM: <http://webserver.dmt.upm.es/~isidoro/Env/Ocean%20thermodynamics.pdf>

Maturana, J., Bello, M., & Manley, M. (2004). Antecedentes históricos y descripción del fenómeno El Niño, Oscilación del Sur. (S. H. Chile, Ed.) *El Niño-La Niña 1997-2000. Sus Efectos en Chile*, 13-27. Recuperado el 04 de junio de 2019, de Divulgameteo: <http://www.divulgameteo.es/uploads/ENSO-antecedentes-descripci%C3%B3n.pdf>

Microsoft. (07 de 05 de 2018). *Librería MDSN*. (Microsoft, Editor, Microsoft, Productor, & Microsoft) Recuperado el 06 de 06 de 2019, de Librería MDSN: <https://docs.microsoft.com/es-es/sql/analysis-services/data-mining/mining-models-analysis-services-data-mining?view=sql-server-2017>

Microsoft. (08 de 01 de 2019). *Microsoft*. Recuperado el 28 de 05 de 2019, de Microsoft: <https://docs.microsoft.com/es-es/sql/analysis-services/data-mining/data-mining-concepts?view=sql-server-2017>

MIRÓ, M. D. (1971). Los foraminíferos planctónicos vivos y sedimentados del margen continental de Venezuela. *Acta Geológica Hispánica*(4), 102-108. Recuperado el 22 de 05 de 2019, de <http://revistes.ub.edu/index.php/ActaGeologica/article/view/3924/4569>

Montealegre-Bocanera, J. E. (31 de 12 de 2007). Modelo institucional del IDEAM sobre el efecto climático de los fenómenos El Niño y La Niña en Colombia. (J. E. Montealegre-Bocanera, Ed.) *Actividades desarrolladas en el marco del contrato de(063-2007)*, 81. Recuperado el 05 de 06 de 2019, de

<http://www.ideam.gov.co/documents/21021/440517/Modelo+Institucional+El+Ni%C3%B1o++La+Ni%C3%B1a.pdf/232c8740-c6ee-4a73-a8f7-17e49c5edda0>

NOAA. (s.f.). NOAA. (C. d. Internet, Productor, & Centro de Predicción del Clima Equipo de Internet) Recuperado el 06 de 06 de 2019, de National Weather Service: https://origin.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ONI_v5.php

Orallo, J. H., & Ramírez, C. F. (2018). Extracción Automática de Conocimiento en Bases de Datos e Ingeniería del Software. *Minería de Datos y Extracción de Conocimiento de Bases de Datos*. Recuperado el 28 de 05 de 2019, de <http://josephorallo.webs.upv.es/docent/doctorat/t2a.pdf>

Pabón, J. (2002). Variabilidad Climática en la América tropical. *Variabilidad Climática en la América tropical*.

Pabón-Caicedo, J. D., & Montealegre-Bocanera, J. E. (1998). *Efectos naturales y socioeconómicos del fenómeno frío del Pacífico [La Niña]*. (IDEAM, Ed.) Recuperado el 05 de 06 de 2019, de IDEAM: <http://documentacion.ideam.gov.co/openbiblio/bvirtual/005192/nina/capitulo2.pdf>

Padilla, M. C. (2010). *Repositorio Institucional de la universidad de Guayaquil*. Recuperado el 22 de 05 de 2019, de Repositorio Institucional de la universidad de Guayaquil: <http://repositorio.ug.edu.ec/handle/redug/807>

Pisico, P. (30 de 06 de 2010). *cambioclimaticoenergia*. Recuperado el 05 de 06 de 2019, de *cambioclimaticoenergia*: <http://cambioclimaticoenergia.blogspot.com/2010/06/la-oscilacion-del-pacifico-tropical-el.html>

- Porras, A. (08 de agosto de 2014). *QantDare*. Recuperado el 29 de julio de 2019, de QantDare: <https://quantdare.com/arboles-clasificacion-matlab/>
- República de El Salvador en la America Central. (2011). ¿Cuáles son los eventos oceanográficos? En R. d. Central, & F. Maquilishuatl (Ed.), *Los eventos oceanográficos* (Vol. 3, págs. 3-6). El Salvador, El Salvador, El Salvador: Fundación Maquilishuatl. Recuperado el 03 de 06 de 2019, de <http://www.marn.gob.sv/wp-content/uploads/los-eventos-oceanograficos-extremos.pdf>
- Ribas, E. (08 de 01 de 2018). *IEBS*. Recuperado el 28 de 05 de 2019, de IEBS: <https://www.iebschool.com/blog/data-mining-mineria-datos-big-data/>
- Riquelme, J., Ruiz, R., & Gilbert, K. (2006). Minería de Datos: Conceptos y Tendencias. *Estado del Arte*, 10(29), 11-17. Recuperado el 06 de 06 de 2019, de <https://idus.us.es/xmlui/bitstream/handle/11441/43290/Miner%C3%ADa%20de%20datos.pdf?sequence=1&isAllowed=y>
- Robles, Y., & Sotolongo, A. (2013). Integración de los algoritmos de minería de datos 1R, PRISM e ID3 a PostgreSQL. (J. J. Inf. Syst., Ed.) *JISTEM - Journal of Information Systems and Technology Management*, 10(2), 389-406. doi:<http://dx.doi.org/10.4301/S1807-17752013000200012>
- Rodríguez, O. (6 de 05 de 2008). Conceptos Básicos. *Conceptos Básicos*. San Pedro, Costa Rica. Recuperado el 06 de 06 de 2019, de http://oldemarrodriguez.com/yahoo_site_admin/assets/docs/Presentaci%C3%B3n_-_Conceptos_B%C3%A1sicos.41132532.pdf
- Rodríguez-Suárez, Y., & Díaz-Amador, A. (Julio-Diciembre de 2009). Herramientas de Minería de Datos. (Y. Rodríguez-Suárez, & A. Díaz-Amador, Edits.) *Red de Revistas Científicas de América Latina y el Caribe, España y Portugal*, 3(3-4), 73-80. Recuperado el 12 de junio de 2019, de <https://www.redalyc.org/html/3783/378343637009/>

- Ruiz, O. (11 de marzo de 2009). *Aplicación de minería de datos para detección de patrones en investigaciones biotecnológicas*. Recuperado el 29 de julio de 2019, de DSpace: <http://www.dspace.espol.edu.ec/handle/123456789/4719>
- Secretaría Nacional de Planificación y Desarrollo. (2017). Plan Nacional Para el Buen Vivir 2017-2021. En S. N. Desarrollo (Ed.), *Planificamos para toda una vida* (pág. 101). Quito: Secretaría Nacional de Planificación y Desarrollo. Recuperado el 30 de 05 de 2019, de <https://www.gobiernoelectronico.gob.ec/wp-content/uploads/downloads/2017/09/Plan-Nacional-para-el-Buen-Vivir-2017-2021.pdf>
- Sinchiguano-Vizuite, M. A. (2017). Manejo Básico De Indicadores De Participación Ciudadana Utilizando Algoritmos Y Técnicas De Minería De Datos Como Apoyo En La Gestión Comunitaria, Caso De Estudio Quitumbe. *Manejo Básico De Indicadores De Participación Ciudadana Utilizando Algoritmos Y Técnicas De Minería De Datos Como Apoyo En La Gestión Comunitaria, Caso De Estudio Quitumbe*. Sangolqui, Pichincha, Ecuador. Recuperado el 11 de Junio de 2019, de <https://repositorio.espe.edu.ec/bitstream/21000/13529/1/T-ESPE-053888.pdf>
- Timarán, S., Hernández, I., Caicedo, S., Hidalgo, A., & Alvarado, J. (2016). El proceso de descubrimiento de conocimiento en bases de datos. (E. U. Colombia, Ed.) *En Descubrimiento*, 63-86. doi:<http://dx.doi.org/10.16925/9789587600490>
- Tobías, A. L., Angulo, A., Daza, A., Toro, D., Gonzalez, J., León, M. V., . . . Quiroga, S. (2014). ZOOPLANCTON. (A. L. Tobías, A. Angulo, A. Daza, D. Toro, J. Gonzalez, M. V. León, . . . S. Quiroga, Edits.) *InfoZOA Boletín de Zoología*, 3, 1-22. Recuperado el 28 de 05 de 2019, de https://www.unimagdalena.edu.co/Content/Public/Docs/Entrada_Facultad3/adjunto_1029-20181004104749_622.pdf

Viana-Pardo, Y. E. (15 de 06 de 2011). *SaberFenomeno*. (Y. E. Viana-Pardo, Editor, Y. E. Viana-Pardo, Productor, & Viana-Pardo, Yesid Elias) Recuperado el 05 de 06 de 2019, de SaberFenomeno: <http://saberfenomeno.blogspot.com/>

Zamorano, E. (28 de 07 de 2018). *El Confidencial*. Recuperado el 31 de 05 de 2019, de El Confidencial: https://www.elconfidencial.com/alma-corazon-vida/2018-07-28/por-que-el-agua-del-oceano-es-salada_1598304/

Zuarth, C. A., Vallarino, A., Jiménez, J. C., & Pfeng, A. M. (2014). *Bioindicadores: Guardianes De Nuestro Futuro Ambiental*. (C. A. Zuarth, A. Vallarino, J. C. Jiménez, & A. M. Pfeng, Edits.) México D.F, México D.F, México: Instituto Nacional de Ecología y Cambio Climático (INECC). Recuperado el 29 de 05 de 2019, de <https://agua.org.mx/wp-content/uploads/2017/11/Bioindicadores-Guardianes-de-nuestro-futuro-ambiental.pdf>

ANEXOS

Tabla 1. Orden y especies procesadas

CLASE	ORDEN
Globothalamea	Calanoida
Gastropoda	Cyclopoida
Ostracoda	Calanoida copepodito
Hexanauplia	Amphipoda
Malacostraca	Aphragmophora
Ophiuroidea	Amphilepidida
Bivalvia	Sessilia
Appendicularia	Salpida
Actinopterygii	Halocyprida
Hydrozoa	Euphausiacea
Polychaeta	Harpacticoida
Branchiopoda	Cheilostomatida
Sagittoidea	Doliolida
Polycystina	Rotaliida
Gymnolaemata	Pesciformes
Scyphozoa	Pteropoda
Thaliacea	Spionida
Invertebrados	Diplostraca
Enteropneusta	Littorinimorpha
Asteroidea	Anthoathecata
Cephalopoda	Myodocopida
Centrohelea	Díptero
Oligotrichea	Copelata
Insecto	Phyllodocida
Nuda	Podocopida
Echinoidea	Leptothecata
Nemátoda	Terebellida
Dinophyceae	Spumellaria
Sipunculidea	Trachymedusae
Leptocardii	Stomatopoda
Tentaculata	Centrohelida
Actinopoda	Caenogastropoda
Lingulata	Choreotrichida
Oligotrichea	Hymenóptera

CLASE	ORDEN
Acantharia	Lingulida

Globothalamea	Cyclopoida copepodito
Phaeodarea	Odonata
Hydrozoa	Noctilucales
Gastropoda	Octopoda
Bivalvia	Isopoda
Hexanauplia	Decapoda (achelata)
Sagittoidea	Golfingiida
Appendicularia	Cydippida
Actinopterygii	Narcomedusae
	Nassellaria
	Siphonophorae
	Sessilida
	Camarodonta
	Rhizostomeae
	Siphonostomatoida
	Beroida
	Pyrosomatida
	phyllodoceomorpha
	Arthracanthida
	Rotaliida
	Phaeocystida
	Siphonophorae
	Pteropoda
	Calanoida
	Harpacticoida
	Sessilia
	Aphragmophora
	Copelata
	Pesciformes
	Coronatae



**Presidencia
de la República
del Ecuador**



**Plan Nacional
de Ciencia, Tecnología,
Innovación y Saberes**



SENESCYT
Secretaría Nacional de Educación Superior,
Ciencia, Tecnología e Innovación

DECLARACIÓN Y AUTORIZACIÓN

Yo, **Macarlupo Cajas, Omayra Violeta** con C.C: # 0926389891 autora del trabajo de titulación: **Modelo predictivo de ocurrencia de eventos extremos de la variabilidad climática en las costas ecuatorianas basado en la detección de patrones de distribución de organismos zooplanctónicos utilizando minería de datos** previo a la obtención del título de **Ingeniera en Sistemas Computacionales** en la Universidad Católica de Santiago de Guayaquil.

1.- Declaro tener pleno conocimiento de la obligación que tienen las instituciones de educación superior, de conformidad con el Artículo 144 de la Ley Orgánica de Educación Superior, de entregar a la SENESCYT en formato digital una copia del referido trabajo de titulación para que sea integrado al Sistema Nacional de Información de la Educación Superior del Ecuador para su difusión pública respetando los derechos de autor.

2.- Autorizo a la SENESCYT a tener una copia del referido trabajo de titulación, con el propósito de generar un repositorio que democratice la información, respetando las políticas de propiedad intelectual vigentes.

Guayaquil, **11 de septiembre de 2019**

f. _____

Nombre: **Macarlupo Cajas, Omayra Violeta**

C.C: **0926389891**



Presidencia
de la República
del Ecuador



Plan Nacional
de Ciencia, Tecnología,
Innovación y Saberes



SENESCYT
Secretaría Nacional de Educación Superior,
Ciencia, Tecnología e Innovación

REPOSITORIO NACIONAL EN CIENCIA Y TECNOLOGÍA

FICHA DE REGISTRO DE TESIS/TRABAJO DE TITULACIÓN

TEMA Y SUBTEMA:	Modelo predictivo de ocurrencia de eventos extremos de la variabilidad climática en las costas ecuatorianas basado en la detección de patrones de distribución de organismos zooplanctónicos utilizando minería de datos.		
AUTOR(ES)	Omayra Violeta Macarlupo Cajas		
REVISOR(ES)/TUTOR(ES)	Galo Enrique Cornejo Gómez		
INSTITUCIÓN:	Universidad Católica de Santiago de Guayaquil		
FACULTAD:	Ingeniería		
CARRERA:	Ingeniería en sistemas computacionales		
TÍTULO OBTENIDO:	Ingeniera en sistemas computacionales		
FECHA DE PUBLICACIÓN:	11 de septiembre del 2019	No. DE PÁGINAS:	98
ÁREAS TEMÁTICAS:	Minería de datos, Variabilidad climática, distribución de organismos zooplanctónicos		
PALABRAS CLAVES/ KEYWORDS:	Minería de datos, Ciclo ENOS, Metodología KDD, Organismos zooplanctónicos, modelo predictivo		
RESUMEN/ABSTRACT:	<p>El presente trabajo tiene como objetivo obtener un modelo capaz de predecir la ocurrencia de eventos extremos climáticos como lo son los del ciclo ENOS (El Niño Oscilación Sur) a partir del comportamiento de los organismos zooplanctónicos, mediante el uso de técnicas de minería de datos, el gran volumen de datos usados pertenecen al programa permanente de Variabilidad Climática del Instituto Nacional de Pesca, quien cuenta con base de datos histórica referente al tema.</p> <p>Las técnicas de minería de datos fueron ejecutadas a través del uso de la metodología KDD (Descubrimiento de Conocimiento en Base de Datos), mediante el uso de la herramienta KMINE, que es una plataforma tecnológica especializada en la ejecución de los algoritmos de minería de datos, tales como: J48 y PRISM, que son algoritmos que pertenecen a la minería de datos de tipo descriptiva, para la cual el tipo de algoritmos que se utilizan son los de clasificación, como los mencionados anteriormente.</p> <p>El modelo predictivo planteado permite determinar la ocurrencia de eventos del Ciclo ENOS tales como: Fenómeno Niño, Fenómeno de la Niña o Condiciones normales, los cuales a través de factores relevantes tomados de la base de datos luego del análisis y proceso de minería, se genera a través de reglas obtenidas por medio de los algoritmos usados.</p>		
ADJUNTO PDF:	<input checked="" type="checkbox"/> SI	<input type="checkbox"/> NO	
CONTACTO CON AUTOR/ES:	Teléfono: +593-4-2086263	E-mail: ovmc94@gmail.com	
CONTACTO CON LA INSTITUCIÓN (COORDINADOR DEL PROCESO UTE)::	Nombre: Toala Quimí, Edison		
	Teléfono: +593-4-220263 ext. 1025		
	E-mail: edison.toala@cu.ucsg.edu.ec		
SECCIÓN PARA USO DE BIBLIOTECA			
Nº. DE REGISTRO (en base a datos):			
Nº. DE CLASIFICACIÓN:			
DIRECCIÓN URL (tesis en la web):			