SELLO EDITORIAL
CORUNIAMERICANA

T.1

TENDENCIAS

DE LA INGENIERÍA WEB Y LA ARQUITECTURA DE SOFTWARE

Autores coordinadores
Roberto Porto Solano
Andres Porto Solano
Diana Suarez Lopez
Adriana María Iglesias Solano

SELLO EDITORIAL
CORUNIAMERICANA

TENDENCIAS DE LA INGENIERÍA WEB Y LA ARQUITECTURA DE SOFTWARE

Compiladores

Roberto Porto Solano
Andres Porto Solano
Diana Suarez Lopez
Adriana María Iglesias Solano

Autores

Diana Suárez López
Roberto Porto Solano
Andrés Porto Solano
Adriana María Iglesias Solano
Roberto Porto Barceló
Luis Sales
Luis Eduardo Trujillo Toscano
Wilmer Acosta Angarita
Hernan Hernandez
Danilo Antonio Garcia Sales
Francisco Hernández
Yesid Rafael Mendoza Cantillo
Martín Emilio Díaz Mora
Yeimer Mora
Miguel Jimeno Paba
Ricardo Antonio Simancas Trujillo
Ricardo Llanos Bravo
Rodrigo Alberto Barbosa Correa
Mauricio Quintana Torres

Libro resultado de investigación, realizado a partir del trabajo colaborativo entre grupos de investigación y el desarrollo de propuestas que contribuyen al fortalecimiento de los indicadores de generación de nuevo conocimiento en el área del Derecho.

Tendencias de la Ingeniería web y la arquitectura de software / Roberto Porto Solano … [et. al.]. -- Barranquilla: Corporación Universitaria Americana, 2018.

Tomo 1; 17x24 cm.
ISBN: 978-958-5512-43-6

1. Ingeniería en la Web -- Tendencias 2. Modelo de ahorro de energía -- Instituciones educativas 3. Bases de datos -- Enseñanza. – Corporación Universitaria Americana. I. Porto Solano, Roberto. II. Porto Barceló, Roberto. III. Sales, Luis. IV. Toscano, Luis. V. Acosta Angarita, Wilmer. VI. Simancas Trujillo, Ricardo. VII. Hernández Belaides, Hernán. VIII. García Sales, Danilo Antonio. IX. Porto Solano, Andrés. X. Hernández, Francisco. XI. Mendoza Cantillo, Yesid Rafael. XII. Cobos Guzmán, Salvador. XIII. Mora, Yeimer XIV. Jimeno Paba, Miguel. XV. Díaz Mora, Martín Emilio. XVI. Llanos Bravo, Ricardo. XVII. Iglesias Solano, Adriana María. XVIII. Barbosa Correa, Rodrigo Alberto. XIX. Quintana Torres, Mauricio.
005.1 T291 2018 cd 21 ed.
Corporación Universitaria Americana-Sistema de Bibliotecas

©Corporación Universitaria Americana
©Sello Editorial Coruniamericana©
ISBN: 978-958-5512-43-6

TENDENCIAS DE LA INGENIERÍA WEB Y LA ARQUITECTURA DE SOFTWARE

Compiladores:
© Roberto Porto Solano, © Andres Porto Solano, © Diana Suarez Lopez
© Adriana María Iglesias Solano

Autores:
© Diana Suárez López, ©Roberto Porto Solano, ©Andrés Porto Solano, © Adriana Iglesias Solano,
© Roberto Porto Barceló, © Luis Sales, © Luis Eduardo Trujillo Toscano, © Wilmer Acosta Angarita,
© Hernan Hernandez, © Danilo Garcia Sales, © Francisco Hernández,© Yesid Mendoza Cantillo,
© Martín Emilio Díaz Mora, © Yeimer Mora, © Miguel Jimeno Paba, © Ricardo Simancas Trujillo,
© Ricardo Llanos Bravo,© Rodrigo Barbosa Correa, © Mauricio Quintana Torres,
© Willys Colpas Peñas, © Salvador Pabón Orosco, © Daladier Jabba Molinares
© Harold A. Pérez Olivera, © Liseth Fontalvo Rueda, © Víctor Higuera Ojito,
© Alex Nader Reales , © José Capacho Portilla

Presidente
JAIME ENRIQUE MUÑOZ
Rectora Nacional
ALBA LUCÍA CORREDOR GÓMEZ
Vicerrector Académico Nacional
MARIBEL YOLANDA MOLINA CORREA
Vicerrector de Investigación Nacional
ASTELIO DE JESÚS SILVERA SARMIENTO
Director Sello Editorial
JUAN CARLOS ROBLEDO FERNÁNDEZ
Sello Editorial Coruniamericana
selloeditorialcoruniamericana@coruniamericana.edu.co

Diagramación y portada: Kelly J. Isaacs González
Corrección de estilo: Eva Luna Contreras Mariño
1a edición: 11 de diciembre de 2018

Todos los derechos reservados. Ninguna parte de esta publicación puede ser reproducida, almacenada en sistema recuperable o transmitida en ninguna forma o por medio electrónico, mecánico, fotocopia, grabación, u otro, sin previa autorización por escrito del Sello Editorial Coruniamericana y de los autores. Los conceptos expresados en este documento son responsabilidad exclusiva de los autores y no necesariamente corresponden con los de la Corporación Universitaria Americana y da cumplimiento al Depósito Legal según lo establecido en la Ley 44 de 1993, los decretos 460 del 16 de marzo de 1995, el 2150 de 1995, el 358 de 2000 y la Ley 1379 de 2010.

Pares Evaluadores
Dougglas Hurtado Carmona
Universidad Metropolitana

Jorge Vengoechea Orozco
Universidad Metropolitana

Comité Científico
Luis Armando Cobo Campo
UNIVERSIDAD EAN

Fabio Moya Camacho.
Universidad de la Guajira.

Pedro Mario Wightman Rojas
Universidad del Norte

Luis Fernando Garces
Universidad Nacional de Colombia

Contenido

Prólogo 7

Modelo de ahorro energético para instituciones educativas MAEIE..... 8
Roberto Porto Solano, Roberto Porto Barceló, Luis Salas, Luis Toscano

Arquitecturas por capas y de servicios web para un catálogo de trámites en línea 28
Roberto Porto Solano, Wilmer Acosta, Hernán Hernández Belaides
Harold Pérez Olivera, Luis Toscano

Aplicación de una estructura conceptual para el desarrollo de sistemas de información en el contexto de ciudades inteligentes..... 43
Danilo García Salas, Wilson Nieto Bernal, Roberto Porto Solano,
Martin Diaz Mora

Metodología para la creación de algoritmos en la educación superior..... 63
Dougglas Hurtado Carmona, Roberto Porto Solano, Israel Escobar Hernandez,
Roberto Porto Barceló, Ricardo Simancas Trujillo

Diseño de un modelo para asistencia médica y predicción de preeclampsia en mujeres embarazadas..... 76
Roberto Porto Solano, Miguel Jimeno Paba, Eduardo Zurék Varela
Francisco Hernández Caballero

Prólogo

La presente obra corresponde a la recopilación de trabajos de grado de maestría y al desarrollo de proyectos financiados por el estado y desarrollados por diferentes universidades. Se ha de resaltar que estos trabajos nacen como resultado de dichos estudios como el resultado de las investigaciones y que cuentan con la dirección y autoría del Doctorando Roberto Porto Solano y la coautoría de docentes, estudiantes de la Corporación Universitaria Americana y la Universidad del Norte los cuales se ven reflejados en cada uno de los escritos que son objetos de esta publicación.

Es igualmente necesario resaltar que para llevar a buen término esta obra, fue el oportuno y eficaz apoyo recibido de la casa editorial de la Corporación Universitaria Americana, de cuyo programa de publicación fuimos beneficiados, y al cual esperamos responder con la altura y responsabilidad de la citada casa de estudios.

El presente libro se encuentra conformado por dos tomos:

En el primer tomo se presenta una contextualización de lo más reciente en proyectos resultados de investigación arraizados en la presentación de nuevas propuestas conceptuales en los temas de inteligencia artificial y sus aplicaciones en el área de la medicina, de arquitecturas orientados a servicios y sus implicaciones en seguridad en los procesos de tramites en línea y la construcción de sistemas para la educación.

En el segundo tomo se presenta una exploración de proyectos resultados de investigación igualmente arraizado en la presentación de nuevos sistemas conceptuales en los temas de sistemas de alertas tempranas para la prevención de accidentes o desastres naturales, adicional mente se trata de análisis de conglomerados para la creación de bases de datos relacionales, por otro lado, se muestran casos de estudio en la industria de los aglomerados productivos locales.

No me cabe duda de que este libro puede establecerse como fuente de consulta para la implementación de soluciones en innovadoras en el mundo real tanto en el contexto académico como en el profesional.

MODELO DE AHORRO ENERGÉTICO PARA INSTITUCIONES EDUCATIVAS MAEIE•

Roberto Porto Solano*
Roberto Porto Barceló **
Luis Salas***
Luis Toscano****

• Capítulo de libro resultado de investigación. Producto desarrollado por el grupo de investigación de Gigetic en colaboración con la empresa Baltrac. Producto resultado del proyecto institucional "Modelo de ahorro energético para instituciones educativas MAEIE" del Politécnico Costa Atlántica (2018).

* Ingeniero de Sistemas. Magister en ingeniería de Sistemas y Computación. Politécnico Costa Atlántica. Correo: rportos@pca.edu.co

** Ingeniero de Sistemas. Magister en ingeniería de Software y Sistemas informáticos. Politécnico Costa Atlántica. Correo: rportob@pca.edu.co

*** Ingeniero eléctrico. Gerente de la empresa Baltrac. Correo: ingeniería@baltrac.com.co

**** Estudiante de Ingeniería de sistemas del Politécnico Costa Atlántico

Resumen

El presente proyecto presenta un modelo que se enmarca sobre el contexto de ciudades inteligentes, para crear una estructura de información para la toma de decisiones a partir de la gestión del servicio público de energía en las instituciones educativas del Distrito de Barranquilla. Dentro del proyecto se implementará un modelo orientado a una arquitectura por capas para el ahorro energético en instituciones educativas. La propuesta de este modelo surge el interés de diseñar e implementar un sistema de tele medición, que permita establecer una comunicación a distancia con un equipo de medición anexo a los contadores de luz, con el fin de obtener la información de lectura de consumo registrada y poder gestionarla mediante un control y seguimiento de indicadores para que sea comparada con datos históricos de consumo energía, con el fin obtener información relevante para la toma de decisiones e implementación de políticas, actividades, tecnologías informáticas que lleven a optimizar el consumo de servicios públicos y minimizando el gasto. El modelo será probado a partir de 672 registros capturados por mes en instituciones del Distrito de Barranquilla arrojando un aumento desmesurado en comparación con datos históricos. Las variables de Watts, VRMS, Amperaje, y Amperaje pico tele medidos también fueron comparados con arrojados en las facturas de la empresa prestadora del servicio de energía, encontrando diferencia sustancial en las mediciones.

Palabras clave: xxx

Abstract

The present project presents a model that is framed in the context of smart cities, to create an information structure for decision making in the management of public energy service in the educational institutions of the District of Barranquilla. Within the project, a model oriented to a layered architecture will be implemented to save energy in educational institutions. The proposal of this model arises the interest to create and implement a system of telemeasurement, which can establish a communication at a distance with a communication equipment attached to the light meters, in order to obtain the consumption reading information Registered and able to manage it by means of a control and monitoring of indicators for the sea compared with energy consumption data, in order to obtain relevant information for the decision making and implementation of policies, activities, computer technologies that lead to optimize the consumption of public services and mini-

mizing spending. The model will be tested from 672 records captured per month in the institutions of the District of Barranquilla. The variables of watts, VRMS, Amperage and Peak Amperage were also compared with the results in the invoices of the utility company, finding the substantial difference in the measurements.

Introducción

Durante el período 2008 – 2011 y dentro del marco del Plan Decenal de Educación aprobado por el Concejo Distrital mediante acuerdo 007 de 2010 y sancionado por el Alcalde Distrital de Barranquilla Alejandro Char Chaljub, la administración Distrital establece la Educación como política pública en el Distrito de Barranquilla, fija las metas en materia de cobertura, calidad y eficiencia de los recursos físicos y administrativos sostenibles, para mejorar la infraestructura física, dotación tecnológica y modernización de los establecimientos educativos, ejecutando durante ese período un gran programa de modernización de la infraestructura física de las instituciones educativas denominado "Plan Alcalde" que adelantó la administración distrital a través de la Secretaría de Educación para modernizar la infraestructura de más de 100 instituciones educativas, y, los proyectos de inversión de Ley 21 de 1982 orientados a garantizar la infraestructura física de las instituciones educativas para hacer posible la continuidad de los alumnos en el sistema educativo en los cuales la Nación a través del MEN aporta el 70% y el Distrito de Barranquilla el 30%, para una inversión total de más de 140 mil millones de pesos en infraestructura física.

La Meta de la administración distrital al finalizar la vigencia 2011 fue que las obras cumplieran con especificaciones que garantizaran espacios dignos, aulas más amplias con aire acondicionado y la dotación necesaria para garantizar una educación de calidad tanto para estudiantes como docentes por lo que es necesaria y urgente la implementación de estrategias y buenas prácticas que permitan un manejo eficiente de los servicios públicos en las instituciones educativas del Distrito de Barranquilla que conlleve a la reducción del consumo y por ende del gasto.

2. Descripción del problema

La ciudad de Barranquilla tiene como propósito la gestión integral de servicios públicos de las instituciones educativas y más particularmente el servicio de energía eléctrica. La problemática alrededor del tema se centra en la utilización ineficiente de los servicios por parte de la comunidad educativa, la indiferencia de algunos hacia lo público, la irresponsabilidad de otros, la falta de coordinación con las empresas de servicios públicos sobre el control del consumo de servicios.

Actualmente las cifras del gasto en servicio público de energía que generan las

instituciones educativas adscritas a la Secretaría de Educación del Distrito son demasiado altas, desde el año 2007 los indicadores de consumo se han disparado en cifras que no son acordes a la realidad de consumo de los entes educativos. Tabla 1. Histórico de consumos de energía en instituciones educativas de Barranquilla.

Tabla 1.
Histórico de consumo de energía en instituciones educativas de Barranquilla

Segmento	Servicios	2007	2008	2010	2011	2012	2013	2014	2015
ESCUELAS Y COLEGIOS	ECA	371.547	186.284	331.421	314.208	369.147	426.046	478.978	527.897

Fuente: Elaboración de los Autores, 2016.

Desde el año 2013 la Secretaría de Educación del Distrito de Barranquilla de manera conjunta con las empresas prestadoras de servicios públicos han modernizado un porcentaje significativo de equipos del sistema de medida convencional al sistema de tele medida, lo cual constituye una oportunidad para la relación empresa – cliente, mejorando por una parte el proceso de lectura y gestión de la facturación para la empresa prestadora del servicio, y, por la otra, sirviendo de insumo a la Administración Pública para la gestión y optimización del consumo de servicios públicos en las instituciones educativas cuyo gasto en los últimos años ha crecido desmesuradamente.

El componente de valor que aporta la implementación esta investigación, va en concordancia con el propósito de la Administración Pública de gestionar integralmente el consumo de servicios en las instituciones educativas (eficiencia energética), la estrategia de la Administración Pública incluye: la optimización del uso de servicios, la reducción de consumos, la reducción de costos y, la implementación de políticas, actividades , tecnologías informáticas, dispositivos y mecanismos que permitan optimizar el uso de recursos y servicios en las escuelas públicas.

3. Métodos de ahorro energético

Se le conoce como ahorro energético a la optimización del consumo energía, cuyo objeto es disminuir el uso de energía, pero produciendo los mismos resultados finales.

Los autores Boudrias y Bannard (2002) proponen que los armónicos en los sistemas de distribución eléctrica se crean a partir de una serie de fuentes, y producen una variedad de efectos secundarios indeseables, por lo tanto, es importante entender todas las soluciones que están disponibles. Proponen la revisión del cambio de fase como un concepto para resolver ciertos tipos de problemas relacionados con la calidad de la energía.

Zebbane, Chenait, y Badache (2012) comentan en su estudio, que el control topológico es la técnica primaria de ahorro de energía y es también uno de los objetivos más importantes El control de topología consiste en mantener un número mínimo de los nodos sensores para operar en modo activo con el fin de ahorrar energía, así como asegurar la conectividad de la red y / o cobertura. En este trabajo proponen un Control de Topología Geográfica (GTC) para reducir el consumo de energía en redes inalámbricas Redes de sensores mediante la programación de estados entre los nodos. El propósito de asegurar la conectividad de la red. GTC divide la red en zonas de modo que una columna vertebral conectada pueda mantener un conjunto necesario de nodos de trabajo y apagando los redundantes. Por lo tanto, asegura la conectividad y prolonga la vida útil de la red seleccionando sólo un nodo activo en cada zona. La simulación del protocolo propuesto supera los resultados en términos de tiempo de vida de la red, entrega de datos, escalabilidad y gastos generales de los mensajes. Puede consumir, respectivamente 80% menos energía que un protocolo de enrutamiento de vida útil de la red.

A comienzos de año, Luo et al, (2017) propusieron que la gestión de la demanda es uno de los temas claves de las redes inteligentes. Su estudio integra el paradigma de computación de servicios en el dominio de la red inteligente y proponen un sistema de recomendaciones personalizadas del lado de la demanda (PRS). El SRP que proponen emplea técnicas de recomendación de servicio para inferir los intereses y necesidades potenciales de los usuarios residenciales en aparatos de ahorro de energía, y luego recomiendan dispositivos de ahorro de energía a los usuarios, creando potencialmente oportunidades para ahorrar energía para la red. El enfoque propuesto se inicia aplicando un método de monitorización de la carga del dispositivo (NILM) no intrusivo basado en el filtrado generalizado de partículas para desagregar los perfiles de utilización de los electrodomésticos de los usuarios finales de los datos del medidor inteligente. Sobre la base de los resultados NILM, se aplican varias reglas de inferencia para inferir las preferencias y patrones de consumo de energía, y para formar el perfil de usuario. En paralelo, se aplican técnicas de recuperación de información para extraer palabras clave de los anuncios de dispositivos de texto y para definir el

perfil del dispositivo. Por último, se aplica el método de medición de similitud para comparar el perfil de usuario y el perfil del dispositivo, clasificar el anuncio del dispositivo y hacer las recomendaciones.

Este mismo año, en su artículo, Rahman, Heerden, y Shafiullah (2016), exponen la importancia del ahorro de energía para reducir costos, y concluyen que el ahorro de combustibles fósiles, las reducciones de emisiones de carbono, así como una caída permanente en los precios de la electricidad, son incentivos importantes para que los consumidores residenciales miren diferentes métodos para reducir su consumo de energía. Su estudio de Respuesta a la Demanda RD es un método alternativo que brinda una oportunidad para que los consumidores reduzcan su costo de consumo de energía al diferir o cambiar su uso de electricidad durante los períodos punta o altos. Con este objetivo, su estudio evalúa la eficacia de las técnicas de RD basada en los precios actualmente disponibles en Australia Occidental, basadas en el costo del consumidor de electricidad y nivel de confort. Las tarifas de electricidad se examinan sistemáticamente y se realiza un análisis de costes adecuado utilizando perfiles de carga realistas de un hogar típico obtenidos mediante el uso de dispositivos inteligentes de monitorización de carga. Se lleva a cabo un enfoque basado en múltiples escenarios para evaluar las tarifas de electricidad para el hogar típico. Los resultados del análisis muestran que la tarifa Smart Home (SM1) sería la elección óptima para el consumidor típico.

Xia y Zhang (2011), expone en su artículo, la necesidad de medir y verificar los ahorros reclamados de los proyectos de eficiencia, si bien las actividades de medición y verificación (M & V) están guiadas por el Protocolo Internacional de Medición y Verificación del Desempeño en muchos países, se cree que el M & V es una ciencia inexacta, es más una práctica de ingeniería con mucho juicio profesional. En su artículo presentan una descripción matemática del problema de M & V y, por lo tanto, arrojan un marco científico de los conceptos, proposiciones, técnicas y metodologías básicas de M & V. Comienza con una descripción general y modelado de sistemas de energía, luego introducen la idea de desarrollar un plan de M & V como un problema de modelado de M & V. Los modelos de M & V se muestran a través de ejemplos que establecen analizando leyes físicas, datos medidos o procesos estocásticos. Como una aplicación de este marco, introducen finalmente una noción de diseño óptimo de M & V.

Estudios recientes han demostrado que el consumo de energía de las redes de acceso inalámbrico es una amenaza para la sostenibilidad de los servicios móviles en la nube. En consecuencia, las soluciones eficientes desde el punto

de vista energético se están volviendo cruciales tanto para las redes de acceso locales como inalámbricas (Amokrane, Langar, Boutaba & Pujolle, 2015).

Las actualizaciones progresivas del stock de edificios existente y la conexión con la ciudad inteligente a través del proceso de modelado de información de construcción (BIM), es un factor fundamental para la implementación de un entorno construido conectado, puesto que hoy en día, este tema central de la Transición Digital, tanto sobre infraestructura (escuelas innovadoras), tanto en los servicios educativos como en los métodos de aprendizaje (Plan Nacional de Metodologías de Escuelas Digitales), logrando un enfoque basado en el rendimiento para el segmento de energía y de acuerdo con una mentalidad de comportamiento (diseño de comportamiento) para la reorganización y definición funcional y espacial (Di Giuda et al. 2016).

Los métodos tradicionales de regulación de las redes de distribución de electricidad presentan varias deficiencias en términos de incentivos, asignación de riesgos, preferencias de los consumidores y valor de la fiabilidad. La tasa de retorno y la regulación basada en el desempeño no ofrecen incentivos tanto para la eficiencia de costos como para la calidad del servicio. Estas estructuras reguladoras también asignan el riesgo de interrupciones casi en su totalidad a los consumidores y no incorporan las preferencias de los consumidores para la fiabilidad / calidad del servicio. Además, las inversiones en la fiabilidad del sistema no se valoran explícitamente. Este documento propone un esquema de seguro para la confiabilidad como una posible solución. El seguro de fiabilidad proporciona incentivos de inversión económicamente eficientes y alivia el riesgo de confiabilidad de los consumidores. Los consumidores proporcionan señales económicas al proveedor de la distribución para su calidad de servicio deseada a través de contratos de seguros. El valor de la fiabilidad para los consumidores se hace transparente, lo que permite a las empresas de distribución (DisCo) tomar decisiones de inversión eficientes. Seguros también asigna el riesgo de interrupción a la DisCo (que controla el sistema), en lugar de los consumidores (que tienen poco o ningún control sobre la fiabilidad). El seguro de fiabilidad desata efectivamente los servicios de entrega y fiabilidad y permite a los consumidores recibir una fiabilidad diferenciada basada en su valor para este servicio permitiendo mejorar tanto la eficiencia como la asignación de riesgos en comparación con las estructuras reguladoras convencionales (Fumagalli, Black, Vogelsang & Ilic, 2004).

Las redes de acceso están evolucionando rápidamente aumentando su capacidad y área de cobertura, haciendo frente a un mayor número de usuarios

y variedad de terminales. Los operadores tienen como objetivo mantener el alto rendimiento de la red y la calidad del servicio, limitando su capital y gastos operativos, por ejemplo, minimizando las inversiones y el consumo de energía mediante el ahorro de energía en los componentes de la red. Para abordar estos desafíos, este trabajo evalúa el consumo de energía, la disponibilidad de conexión y el tiempo de detección de fallos de tres esquemas de protección aplicables a redes de acceso convergentes: protección de fibra disjunta, protección de fibra disjunta energéticamente eficiente y protección de fibra disyuntiva reflectante. Los esquemas se evalúan mediante un estudio de caso que considera un sistema público inteligente de transporte (ITS) (Grigoreva, Wong, Furdek, Wosinska, & Mas Machuca, 2017).

Otros autores como Ríos, Arévalo, Causarano (2016), explican cómo mejorar la calidad y reducir el consumo del servicio eléctrico de un área metropolitana mediante un análisis de los indicadores de continuidad del suministro. Como finalidad, se propone platicar las inspecciones visuales de líneas aéreas de media tensión (23 KV), basándose en el desempeño del Sistema Metropolitano de Distribución de la Administración Nacional de Electricidad (ANDE), además de estandarizar los niveles de calidad deseados a través de metas sobre distintos plazos, cuyo cumplimiento está supeditado a la incorporación de cuadrillas exclusivas de mantenimiento preventivo.

Solano et al. (2017) proponen un marco de trabajo ágil, orientando en una arquitectura por capas en la que implementan un sistema para el control de recetas médicas utilizando la tecnología de comunicación de campo cercano. En el cual el paciente y medico utilizan una aplicación móvil en que permiten la creación de recetas médicas, la visualización y auto renovación para hacer reclamación de medicamentos en las entidades prestadoras de servicios de salud. Nuevamente Solano et al. (2017) hacen un planteamiento de un marco de enseñanza orientado a una arquitectura por 4 capas, para el adiestramiento en la introducción de los algoritmos basado en objetos virtuales de aprendizaje llamado OLIA.

4. Sistemas de tele-medición

Para Antonyuk, Varshavskiy, Kolpakova, y Minina (2016), los sistemas de tele medición y procesamiento de datos está aumentando, por lo que es necesario determinar y / o controlar las grandes cantidades de información por características para examinar diferente objetos. Él explica que cada uno de los

datos es diferente y que no todos los datos que se reciben en la medición son importantes.

Por otro lado Aiad (2017), hace un monitoreo de carga no intrusiva que analiza mediciones agregadas totales de medidores inteligentes. A través de este monitoreo pueden informar a los propietarios de viviendas acerca de consumos detallados por dispositivos individuales, con el fin de poder influir en el comportamiento de consumo de energía, lo que resulta un ahorro de energía.

Según Damjanovic, Integlia, y Sarwat (2016), explican que los sistemas de medición se consideran como una red de dos puertos en la que las pérdidas se disipan. Ellos propusieron varias conexiones para mediciones de pérdidas de transformadores de potencia en condiciones no lineales utilizando instrumentos digitales y diferentes métodos de medición: Diferencia de Entrada y Salida de Potencias, denominada Método de Entrada y Salida de Potencia, y Diferencia de Voltaje y Corriente. El tema de este artículo es discutir y evaluar la precisión en las mediciones de las pérdidas del transformador con esos métodos.

Otros como Amberg, Gepp, Kozik, Foehr, y Vollmar (2014), proponen que los métodos de medición de rendimiento en las organizaciones de ingeniería son importantes para diseñar, implementar y controlar las actividades de mejora, por lo cual los métodos de medición del rendimiento son utilizados en la práctica industrial. Este enfoque se centra en los métodos de evaluación de inversiones (como la mejora de la eficiencia de programas) en los departamentos de ingeniería basándose en un estudio empírico a través de una encuesta a 56 personas en proyectos de I+D.

En otro trabajo Koprivica, Member, y Nešković (2016), proponen la eficiencia y mejora de los métodos de medición del campo eléctrico en las proximidades de las estaciones base, utilizando un procedimiento para el uso de intervalos de tiempo promediados (en comparación con los 6 minutos que estipulan las normas) y un modelo para la conversión de medidas monoaxiales a Isotrópico.

La medición de cerca del ruido de campo eléctrico de bombillas LED mediante el uso de antena o monopolo vertical, obtienen el factor de calibración por un método de antena de campo cercano tres, utilizando alternativamente dos antenas de longitud diferente, de 1 a 400 MHz, logrando observar el campo eléctrico cercano característico de 100 a 200 Megahercio por lo que se ajustó al campo eléctrico lejos de los bulbos del LED, y se ha correlacionado con el campo eléctrico cercano (Kasuga, Ohashi, Yamada & Inoue, 2012).

5. Metodología

Este modelo está definido dentro de un tipo de investigación analítica-aplicada, por lo que se realizaron múltiples comparaciones de datos históricos de 10 años de consumo energético, además de diseñar e implementar un sistema de tele medición, que permitió establecer una comunicación a distancia con un equipo de medición anexo a los contadores de luz, con la que se obtuvo información relevante para la toma de decisiones e implementación de políticas que permitieron optimizar el consumo de servicios público de energía y minimizar el gasto. El método de investigación es inductivo puesto que a partir de una problemática particular como lo es el del desperdicio energético, se piensa llegar a una solución general como lo es una plataforma basada en un modelo de ahorro energético. Tiene un enfoque cuantitativo, puesto que para diseñar el modelo de ahorro energético es necesario obtener datos por lo que se realizaron capturas de 672 registros capturados por mes en una institución del Distrito de Barranquilla arrojando un aumento desmesurado en comparación con datos históricos. Las variables como Watts, VRMS, Amperaje, Amperaje pico tele medidos, fueron comparados con los arrojados en las facturas de la empresa prestadora del servicio de energía, encontrando diferencia sustancial en las mediciones, estas variables son necesarias también para obtener el indicador de desempeño energético IDEn y compáralo con la línea de base energética LBEn histórica que fueron fundamentales para la definición de métricas necesarias para la determinación de la eficiencia del modelo.

Los pasos metodológicos se describen a continuación.

- Hacer un recorrido del estado del arte sobre técnicas usadas comúnmente para ahorro energético, y los criterios útiles para comparar cuáles técnicas son las mejores, como precisión de la medición, estrategias de mejora, perfil de carga recibido, optimización de la eficiencia, usos de fuentes renovables, facilidad de implementación en cuestiones de costos etc.

- Aprobación del proyecto por parte de la secretaria de educación de la ciudad de Barranquilla.

- Autorización por porte de la empresa Electricaribe S.A para la instalación de equipos medidores y captura de variables.

- Determinar el grupo de variables adecuado para la medición a partir de las normas ISO 50001.

- Diseñar e implementar una plataforma que permite llevar el control de las variables definidas por las normas ISO 5001, para ser tomadas en cada institución.

- Estudio sistemático de la información obtenida.

- Definir unas métricas de ahorro basada en consumos históricos por institución.

- Realizar planes de optimización del consumo energía a partir la selección de dispositivos de ahorro energético.

- Definir nuevas mediciones a partir de métricas para comparar con datos históricos en la plataforma de control

- Plan de instalación de equipos ahorradores de energía.

6. Modelo de ahorro energético MAEIE

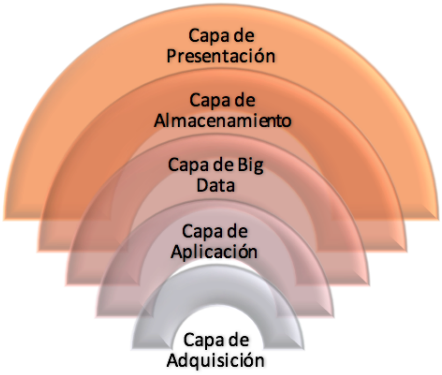


Figura 1. *Modelo de Capas MAEIE.*
Fuente: *Elaboración propia, 2017*

El modelo propone 5 capas, las cuales son fundamentales para desarrollar un sistema de información y de pronósticos de consumo en el entorno de Emerging Cites como lo muestra en la Figura 1. MAEIE propone el análisis de un LBE, línea base energética histórica con la línea base

energética actual, con el fin de ayudar a crear un conjunto de políticas de consumos y planes de selección de equipos en las instituciones educativas. Una particularidad del modelo es que una de las capas es transversal a otras permitiendo menor cohesión mayor acoplamiento del modelo. Seguidamente estaremos describiendo cada aspecto fundamental del modelo, así como cada uno de los detalles y características de las capas que lo componen:

6.1. Capa de Adquisición e infraestructura ubicua

Esta capa hace referencia a la obtención de la información a través de dispositivos ubicuos. En este modelo, los dispositivos ubicuos son utilizados encima de cada contador de energía de las instituciones actuando como módulo de comunicación GSM/GPRS, proporcionando un puente entre la capa de adquisición y la capa de aplicaciones. Este dispositivo permite incrementar las prestaciones del contador de energía, facilitando la interacción con el sistema de información, posibilitando el acceso a toda la información correspondiente al consumo de energía durante periodos de tiempo definidos. Todos los módulos de comunicación deben ser configurados dentro del sistema de tele medición de consumos.

6.2. Capa de Aplicaciones

La capa de aplicación es la encargada de que todo funcione por lo tal no siempre es visible para el usuario y en ella se encuentran las aplicaciones cuya lógica programada son usadas dentro de los terminales finales para optimizar recursos por parte de los usuarios de tipo administrador, o aquellas aplicaciones para administrar las bodegas de datos, en este caso, la conexión es pertinente realizarla hacía la bodega de datos. La capa de aplicaciones se conecta con la capa de adquisición de datos por medio del protocolo IP a la tarjeta GSM/GPRS y esta a su vez de manera serial se conecta al dispositivo ubicuo permitiendo desde la plataforma web por medio de instrucciones AT controlar, administrar, especificar las variables a auditar. Esta a su vez administra las bodegas de datos

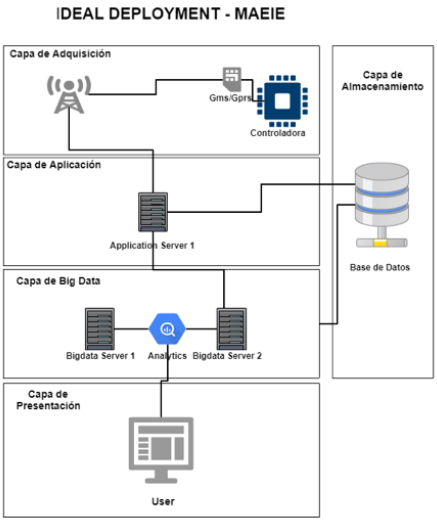


Figura 2. *Despliegue del Sistema.*
Fuente: *Elaboración propia, 2017*

6.3. Capa de Big data

Esta capa es la encargada de hacer un diagnóstico de consumo por institución a partir de los datos arrojados por la capa de aplicación y para los cuales son consultados los datos en la capa de almacenamiento, para que posteriormente se realice un análisis de los datos históricos de consumo de las instituciones educativas en contraste con los datos de indicadores de desempeño energético con el fin de definir un plan de acción de ahorro para cada institución a partir de políticas de consumo.

6.4. Capa de Almacenamiento

En esta capa es en donde reside la información, también es la encargada de recibir todas las solicitudes de almacenamiento y de recuperación de información en la BD bases de datos provenientes de la capa de aplicaciones, por lo que no se casa en una tecnología de BD en específico, bien pueden ser bases de datos relacionales o bien no relacionales. Aunque muy distinta a la capa de adquisición, aplicaciones y big data esta capa debe garantizar disponibilidad, puesto que es trasversal a las capas mencionadas anteriormente.

6.5. Capa de Presentación

Esta capa es la encargada de presentar y capturar la información del

usuario en un mínimo de proceso. También es conocida como interfaz gráfica (frontend) y debe tener la característica de ser entendible para el usuario por resumidos elementos visuales o gráficos, también se realizan las validaciones de los datos que se ingresar al sistema. Esta capa está en constante comunicación únicamente con la capa de aplicaciones la cual maneja la lógica del sistema.

7. Resultado

Dentro de los resultados de esta investigación, tenemos 2 componentes, el modelo de ahorro energético para instituciones educativas expuesto anteriormente y el núcleo principal del sistema, el software Energix. El núcleo principal del sistema es utilizado, a través de dispositivos del Internet de las Cosas en manos de los directores de las instituciones, quienes son los encargados de administrar de forma eficiente los servicios públicos brindados por el estado. El modelo de aprendizaje MAEIE es la parte más importante del modelo de ahorro energético, el cual gran parte está inmerso en la plataforma web, puesto que es la encargada de insertar y mostrar los datos tele medidos de los controladores además de que permite realizar el diagnostico de los valores obtenidos por mes con los datos históricos y definir cuando una institución debe ser auditada por parte de secretaria de educación para aplicar medidas de ahorro dentro de las que se encuentran los planes de selección de equipos eficientes y mantenimientos de equipos definidas en la ISO 5001.

La plataforma web está compuesto por el módulo de lecturas, módulo de diagnóstico y el módulo de presentación.

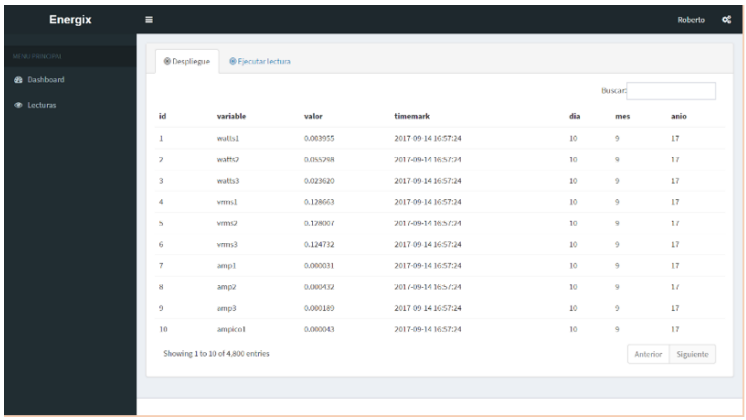


Figura 3. *Módulo de lecturas*
Fuente: *Elaboración propia, 2017*

En el módulo de lectura se muestran las variables watts, vrms, amperaje y amperaje pico. Esta se encuentra sectorizadas por zonas que van desde la zona 1 a la zona 3 juntos con los tiempos de capturas. Las lecturas se realizan cada hora, por lo que se realizan 24 por días con el fin de calcular el promedio por día de cada variable, para posteriormente hallar el promedio mensual. Cabe resaltar que estas lecturas al módulo GSM del controlador a través de comandos AT como se muestra en la Fig. 3.

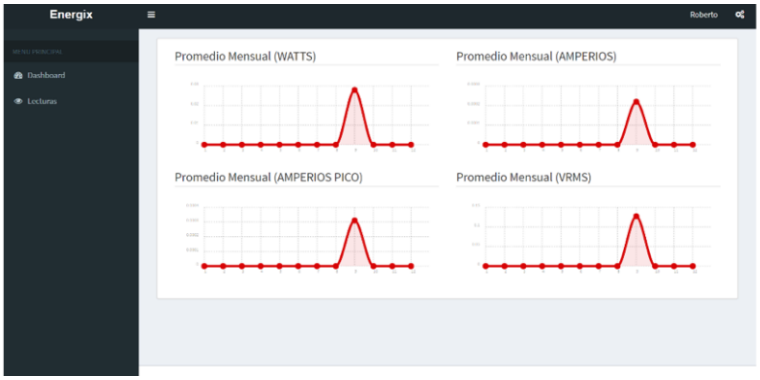


Figura 4. *Módulo de Presentación de Consumo*
Fuente: *Elaboración propia, 2017*

El módulo de presentación muestra el consumo mensual de las diferentes variables descritas anteriormente de cada institución con el fin de poder ver diferencias de consumo de los meses en los que se ha realizado la captura.

8. Discusión

La finalidad de este modelo, como fue mencionado anteriormente, es el auditar el consumo de energía de las instituciones educativas de la ciudad de Barranquilla con el fin de aplicar políticas de ahorro basadas en la norma ISO 5001 a partir de modelo de capas MAEIE definido anteriormente. Más, sin embargo, el controlador no solo permite enviar las lecturas del consumo de energía sino también el verificar la calidad de la fuente de energía del proveedor de servicio de energía. Por lo que se encontró que el proveedor no estaba suministrando lo contratado por la institución, más sin embargo estaba cobrando lo contratado.

9. Conclusión

Esta investigación pretendió el adaptar el ahorro de energía a las políticas de ahorro de energía de forma inteligente esto se logra viendo el uso de los equipos reales y aplicando no solo políticas estáticas en horarios fijos. Los resultados mostraron que podemos lograr ahorros importantes y consideramos que nuestra propuesta puede aplicarse fácilmente en cualquier institución Educativa y cualquier Contador Eléctrico no Inteligente. Puede haber algunas preocupaciones, que pensamos nuestra arquitectura maneja, pero no podemos comprobar por el momento.

También se determinó, que gran parte de los aumentos de los consumos energético en las instituciones educativas proviene de la mala medición y malos cálculos que realiza el proveedor de servicio de energía, puesto que presta un servicio deficiente y por debajo de lo contratado por el usuario. Además, se determinó que, en las instituciones, que se encuentran en las poblaciones con mayor desigualdad social, se encontraron en horarios de baja concurrencia de población estudiantil, se presenta gran consumo de energía, por lo que se comprobó que el gasto viene de las viviendas alrededor de las instituciones.

Referencias Bibliográficas

Aiad, M. (2017). Non-intrusive Monitoring of Overlapping Home Appliances Using Smart Meter Measurements. Power and Energy Conference at Illinois, (February). Retrieved from https://www.researchgate.net/profile/Misbah_Aiad/publication/313916035_Non-intrusive_Monitoring_of_Overlapping_Home_Appliances_Using_Smart_Meter_Measurements/links/58afc6ae92851cf7ae890487/Non-intrusive-Monitoring-of-Overlapping-Home-Appliances-Using-Smar.

Amberg, M., Gepp, M., Kozik, T., Foehr, M. & Vollmar, J. (2014). Use of performance measurement methods in engineering organizations—An empirical study. In Industrial Technology (ICIT), 2014 IEEE International Conference on (pp. 919-924). IEEE.

Amokrane, A., Langar, R., Boutaba, R. & Pujolle, G. (2015). Flow-Based Management for Energy Efficient Campus Networks. IEEE Transactions on Network and Service Management, 12(4), 565–579. https ://doi.org/10.1109/TNSM.2015.2501398

Antonyuk, E. M., Varshavskiy, I. E., Kolpakova, I. S. & Minina, A. (2016). Telemetry System with Adaptive Commutation, 389–391.

Boudrias, J. G. & Bannard, J. (2002). Power factor correction and energy saving with proper transformer and phase shifting techniques. In Petroleum and Chemical Industry Conference, 2002. Industry Applications Society 49th Annual (pp. 99-102). IEEE.

Damjanovic, A., Integlia, R., & Sarwat, A. (2016). Evaluation of power transformer loses measurements methods under nonlinear load conditions. In Industrial and Commercial Power Systems Technical Conference (I&CPS), 2016 IEEE/IAS 52nd (pp. 1-5). IEEE.

Di Giuda, G. M., Villa, V., Piantanida, P., Tagliabue, L. C., Rinaldi, S., De Angelis, E. & Ciribini, A. L. C. (2016). Progressive Energy Retrofit for the educational building stock in a Smart City. IEEE 2nd International Smart Cities Conference: Improving the Citizens Quality of Life, ISC2 2016 - Proceedings, (1), 3–8. https://doi.org/10.1109/ISC2.2016.7580818.

Fumagalli, E., Black, J. W., Vogelsang, I. & Ilic, M. (2004). Quality of service

provision in electric power distribution systems through reliability insurance. IEEE Transactions on Power Systems, 19(3), 1286–1293. https://doi.org/10.1109/TPWRS.2004.831294

Grigoreva, E., Wong, E., Furdek, M., Wosinska, L. & Mas Machuca, C. (2017). Energy Consumption and Reliability Performance of Survivable Passive Optical Converged Networks: Public ITS Case Study. Journal of Optical Communications and Networking, 9(4), C98. https://doi.org/10.1364/JOCN.9.000C98

Kasuga, T., Ohashi, T., Yamada, S., & Inoue, H. (2012). A study on measurement method of near electric field noise from LED bulbs. In SICE Annual Conference (SICE), 2012 Proceedings of (pp. 428-432). IEEE.

Koprivica, M., Nešković, N., & Nešković, A. (2016, November). Improving the efficiency of methods for measurement of electric field strength in the vicinity of public mobile system base stations. In Telecommunications Forum (TELFOR), 2016 24th (pp. 1-1). IEEE.

Luo, F., Ranzi, G., Kong, W., Dong, Z. Y., Wang, S. & Zhao, J. (2017). Non-intrusive energy saving appliance recommender system for smart grid residential users. IET Generation, Transmission & Distribution, 11(7), 1786-1793.

Rahman, M. M., Van Heerden, C. & Shafiullah, G. M. (2016). Electricity Tariffs Evaluation Using Smart Load Monitoring Devices for Residential Consumer. In Computer Science and Engineering (APWC on CSE), 2016 3rd Asia-Pacific World Congress on (pp. 265-270). IEEE.

Ríos, D., Arévalo, M., Causarano, W. (2016). Electrical Service Quality Improvement in Asunci ´ on and Metropolitan Area through an Analysis of Reliability Indexes.

Solano, R. P., Barceló, R. P., Gómez, A. C., Barbosa, J. C., Gutiérrez, C. E., De los Ríos, J. & Meza, E. H. (2017). Framework ágil para el control de recetas médicas que utiliza la tecnología NFC (FARM). Revista Lasallista de Investigación, 14(1).

Solano, R. P., Sarmiento, A. S., Giraldo, L. F., Solano, A. P., & López, D. S. (2017). MARCO DE ENSEÑANZA BASADO EN OVAs PARA EL ADIESTRAMIENTO EN LA INTRODUCCIÓN DE LOS ALGORITMOS

(OLIA). Revista científica electrónica de Educación y Comunicación en la Sociedad Del Conocimiento, 17(1).

Xia, X. & Zhang, J. (2011, September). Mathematical description of the performance measurement and verification. In AFRICON, 2011(pp. 1-6). IEEE.

Zebbane, B., Chenait, M. & Badache, N. (2012). Energy-efficient protocol-based sleep-scheduling for wireless sensor networks. In Complex Systems (ICCS), 2012 International Conference on (pp. 1-6). IEEE.

ARQUITECTURAS POR CAPAS Y DE SERVICIOS WEB PARA UN CATÁLOGO DE TRÁMITES EN LÍNEA•

Roberto Porto Solano*
Wilmer Acosta**
Hernán Hernández Belaides ***
Harold Pérez Olivera****
Luis Toscano*****

• Capítulo de libro resultado de investigación. Producto Colaborativo de los grupos de investigación Gigetic, del Politécnico Costa Atlántica con Aglaia y Procont con la Corporación Universitaria Americana. Producto resultado del proyecto institucional Pertinencia e impacto social de la americana (2016).

* Ingeniero de Sistemas. Magister en ingeniería de Sistemas y Computación. Politécnico Costa Atlántica. Correo: rportos@pca.edu.co
** Ingeniero de Sistemas. Magister en ingeniería de Sistemas y Computación.
*** Contador Público. Magister universitario en fiscalidad internacional. Corporación Universitaria Americana.
**** Ingeniero Industrial. Magister en ingeniería Industrial. Corporación Universitaria Americana.
***** Estudiante de Ingeniería de sistemas del Politécnico Costa Atlántico.

Resumen

El mundo de los servicios web está evolucionando rápidamente. En los últimos cinco años, las especificaciones del servicio web se han definido, refinado y, a veces, se ha criticado por los lanzamientos de nuevos kits de herramientas que facilitan a los desarrolladores diferentes servicios web que son utilizados para construir sistemas, por lo que los fabricantes han trabajado para garantizar la interoperabilidad de estos. Al igual que con la infraestructura web clásica, estos desarrollos e implementaciones tecnologías de servicios web en paralelo, han avanzado a través de un ciclo de retroalimentación positiva. Si bien los servicios web han avanzado mucho en los últimos tres años, el trabajo continúa en la adaptación específica de cada plataforma. Por lo que esta investigación va encaminada en utilización de un marco de trabajo que permita la adaptabilidad de diferentes servicios web sin perder interoperabilidad en las comunicaciones de trámites en línea, viendo en donde se encuentra la modelo de plataformas actuales y el del futuro.

Palabras clave: Marco de trabajo, servicio web, trámites, interoperabilidad, adaptabilidad.

Abstract

The world of web services is evolving rapidly. In the last five years, the specifications of the web service have been defined, refined and, at times, criticized for the launching of new toolkits that provide developers with different web services that are used to build systems, so that the manufacturers have worked to guarantee the interoperability of these. As with the classic web infrastructure, these developments and implementations parallel web services technologies have advanced through a positive feedback loop. Although web services have advanced a lot in the last three years, the work continues in the specific adaptation of each platform. So, this research is aimed at using a framework that allows the adaptability of different web services without losing interoperability in online paperwork communications, seeing where the current and future platforms model is located.

Key words: Framework, web service, procedures, interoperability, adaptability

Introducción

Las TIC se han convertido en una herramienta para el fortalecimiento de los procesos de participación ciudadana, y son estos la razón de ser de los gobiernos. El 30 de Julio de 2009 se creó bajo la Ley 1341 el Ministerio de Tecnologías de la información y las comunicaciones MinTic, donde uno de sus principales derroteros es cerrar la brecha que existe entre el estrado y los ciudadanos; en pro de dar solución a esta necesidad se han propuesto diferentes estrategias y herramientas.

En un mundo digital como el de hoy donde los países desarrollados cuentan con índices de penetración de las TIC y donde la relación con el estado sin las TIC´s es inconcebible.

En Colombia se han desarrollado iniciativas en busca de proporcionar a los ciudadanos un catálogo de trámites y servicios en línea universal a partir de la adopción de arquitecturas orientadas a servicios web en el desarrollo de plataformas, buscando acotar la brecha digital existente y proporcionar a los ciudadanos nuevos medios de participación con el estado.

A pesar de que los Servicios Web (Web Service), son una tecnología que han alcanzado cierto nivel de madurez teórico más que a nivel práctico, siguen en desarrollo por lo que aún falta mucho por definir y realizar. Los protocolos y estándares son una tecnología que aún se está desarrollando, por lo que hay muchas cosas que aún no están del todo claras por la adopción de gran cantidad de estándares abierto para mejorar la interoperabilidad. Los protocolos han aumentado su utilización actualmente a partir del seguimiento estándares que permiten mejorar la calidad de los servicios mas no aseguran el cien por ciento de calidad.

2. Trabajos relacionados

Existen diferentes propuestas para dar solución a trámites en línea, a continuación se referencian algunos de los más importantes:

Mohammadi, Nematbakhsh, y Nematbakhsh (2012), explican las iniciativas de un diseño de un tramitador en línea, que se resume como un sistema que permite administrar un conjunto de distintos servicios para la permutación de información que hacen parte de los trámites y servicios ofrecidos por la organización nacional,

por medio una SOA (Arquitectura Orientada a Servicios) y empleando lenguajes para la permutación de la información, favoreciendo la unificación y optimización de procesos.

Los TL (Tramitadores en line) hacen empleo de tecnología estándar como los Web Services, Protocolo Simple de Acceso a datos, Descripción del Lenguaje de Web Services y Seguridad de Web Services, posibilitando que cualquier proveedor pueda desarrollar nuevos servicios (Kennedy, Stewart, Jacob & Molloy, 2011).

WSDL (Web Services Description Languaje) es un estándar que indica de qué forma debe definir la localización del servicio, los parámetros y cómo se devuelve el resultado. La especificación de WSDL es una recomendación de la W3C y está basado en XML. Cuando un proveedor desea publicar un servicio, debe registrar el WSDL correspondiente al mismo usando UDDI, lo cual permite a las aplicaciones que desean consumir dicho servicio obtener toda la información necesaria (Lessen, Nitzsche & Leymann, 2009).

Una vez ubicado el servicio, la comunicación entre el consumidor y el proveedor se realiza utilizando el protocolo SOAP (Simple Object Access Protocol) que está basado en XML y utiliza HTTP como medio de transporte. Este protocolo también es un estándar recomendado por la W3C que por estar basado en XML es independiente de la plataforma que se elija y del lenguaje de programación (Pinzon, Bajo, De Paz & Corchado, 2011).

La Ciudad de México cuenta con un catálogo de trámites y servicios en línea gubernamentales llamado CDMX, expuesto a los ciudadanos donde muchos de estos tienen un flujo automático de verificación, lo que permite disponer de una ventanilla electrónica 24 X 7 (Gobierno Ciudad de México, 2017).

La adaptación de un motor BPM el cual incluya la integración nativa con diferentes tipos de servicios web, se convertirá en el principal modo de buscar la información.

Los trámites y servicios son tan variables como la tecnología y más cuando hablamos del sector público, por eso el esfuerzo del Ministerio de Tecnologías de la Información y las Comunicaciones en estandarizar una plataforma que sea duradera son loables pero carece de un componente valioso en este proceso "Flexibilidad", sumando a esto la admirable tarea de intentar de crear un estándar para interconectar todos los sistemas de información, labor que carece de durabilidad cuando no se cuenta con el respaldo de la comunidad en la web. Los

entes gubernamentales carecen de la continuidad en la inversión de Investigación y Desarrollo disminuyendo así la posibilidad de perfeccionar los estándares propuestos y dejando que estos queden solo como caso de estudio.

A este punto se hace necesario contar con la integración de diferentes componentes de software bastante pequeños y genéricos de tal modo que reemplazar alguno, genere el menor trauma posible para la arquitectura planteada.

Basados en un motor BPM de código libre, la Alcaldía de Barranquilla en convenio con el MinTic desarrollaron una herramienta que permite modelar diferentes trámites y servicios la cual puede conectarse a través de servicios web a las bases de datos y sistemas legacy de cada entidad territorial.

Se implementó JBPM, el cual es un motor flexible para modelar procesos de negocio, para el desarrollo se han tomado las clases del motor JBPM y se han extendido dando funcionalidades muy propias de los trámites y servicios con las entidades de la república de Colombia. Cuenta con las siguientes características a resaltar:

- Basado en el IDE Eclipse.
- Editor web que soporta la creación de procesos de negocio usando drag & drop.
- Adaptación nativa con JPA / JTA
- Integración con framework de desarrollo en Java
- Administración de listas de tareas
- Monitoreo por consola de los diferentes componentes que intervienen en el proceso
- Conexión con bases de datos remotas y sistemas legacy a través de web services

Con estas características se puede asegurar la flexibilidad en los trámites y servicios según lo demanden los usuarios o la ley colombiana, se utilizan piezas de software soportadas por la comunidad la cual nos permite contar no actualizaciones, se diseña la solución de tal manera que los componentes de software son fácilmente removibles asegurando la rápida adaptación a los cambios acelerados de tecnología a los que nos enfrentamos hoy en día (Porto et al, 2017).

La seguridad de la transmisión de la información se constituye en una pieza fundamental dentro de la construcción de la arquitectura propuesta enviando la información cifrada, basados en RSA algoritmo asimétrico que cifra tramas, que utiliza una llave única tipo pública, la cual es puesta a disposición junto con otra

privada la cual es almacenada de forma confidencial por el titular. Adicional para proteger los paquetes en la red y que estos no puedan ser clonados se implementan certificados digitales, sumándole un usuario y contraseña en la cabecera del mensaje, blindando así la confidencialidad de la información y la persistencia de esta.

La interconexión de la plataforma de trámites y servicios desarrollada por el ministerio con las bases de datos locales de cada entidad se realizan a través de unos servicios web, la solución para el consumo de WS desde la herramienta transaccional consta de varios componentes los cuales se pueden dividir en 2 roles, unos para el desarrollador del proceso y otros para el desarrollador del Web Service (Porto, 2016).

3. Proceso para la creación de trámites o servicios

La publicación de un nuevo trámite o servicio por parte de una Entidad en la Plataforma de Interoperabilidad sigue el siguiente flujo de actividades como se muestra en la Fig. 1 Diseño del flujo de ejecución del trámite o servicio:

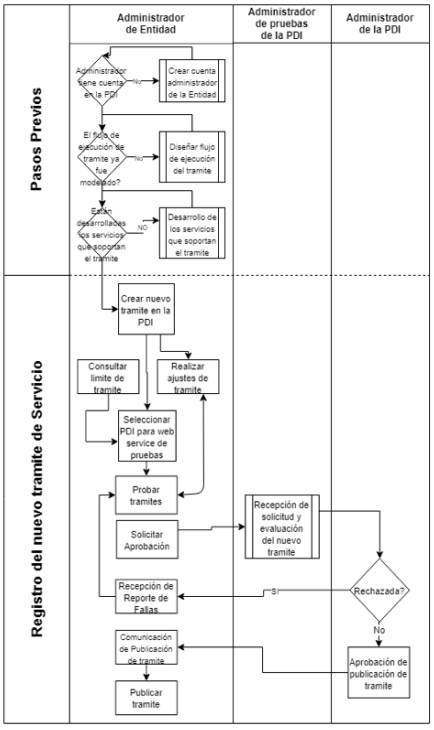


Figura 1. *Diseño del flujo de ejecución del trámite o servicio.*
Fuente: *Elaboración propia*

La Interoperabilidad, éste debe ser primero modelado. Para modelar el trámite o servicio debe seguirse el proceso descrito en el estándar GEL-XML, tanto para la especificación de datos de entrada como la especificación de datos de salida. Los detalles de las tareas sugeridas para modelar el trámite o servicio y generar los esquemas GEL-XML están relacionadas en los documentos del estándar GEL-XML. Los resultados finales de este proceso son:

- Un esquema XSD de entrada y
- Un esquema XSD de salida.

Ambos esquemas existen como archivos (.XSD) y/o almacenados en un Repositorio de Esquemas GEL-XML con acceso desde la PDI.

3.1. Desarrollo de los servicios de soporte al trámite o servicio

Los trámites o servicios de la entidad que podrán ser habilitados para conectar desde la Plataforma de interoperabilidad deben residir en un Web Service con características específicas. Estas características requeridas serán descritas en el curso de este capítulo y calificarán el cumplimiento (compliant) del Web Service con la Plataforma de interoperabilidad. Cuando un Web Service satisface estos requerimientos y se califica como compliant, el Web Service se convierte en un Adaptador de Agencia Conectada – AAC – para la Plataforma de interoperabilidad.

Los Web Services son una tecnología esencial para el crecimiento de la computación distribuida por medio de redes compartidas. Estas se han convertido en un gran medio para la integración de gran cantidad aplicaciones, además de estándares o procesos libre y al realce en la comunicación y contribución entre organizaciones y manejos. Las aplicaciones se estructuran integrando múltiples tecnologías de servicios web de ascendencia distinta, que se desempeñan simultáneamente, no importando su localización o el cómo de implemento (Pautasso, 2009).

3.2. La plataforma de Interoperabilidad

Los servicios web entregan a los usuarios gran cantidad de utilidades a partir de un protocolo estándar en sus plataformas para interoperar como lo es SOAP. Estos servicios web o Web Service ayudan a dar en detalle el funcionamiento de sus interfaces para otros usuarios desarrollen sus aplicaciones clientes para establecer comunicación con el Web Service. Las descripciones proporcionan una estructura base de un documento de la Descripción del Lenguaje de

los Servicios Web (WSDL). Los conceptos acerca de los servicios web y sus utilidades aumentan exponencialmente por su excelente rendimiento en la integración de plataformas e interoperabilidad con otros programas en diferentes lenguajes y plataformas convirtiendo en su más grande ventaja, la comunicación estándar.

SOAP es el protocolo de comunicación estándar menos complejo que otros, por lo que resulta más fácil que una implementación cumpla con los estándares de SOAP. Todos los Web Services se basan en protocolos estándar, como XML, HTTP y TCP/IP. Muchas entidades gubernamentales y empresas que tienen una infraestructura Web, así como de personal y conocimientos para mantenerla, de tal forma que el esfuerzo y costo de entrada a estos servicios es mucho menor que con otras tecnologías. Por esto, las plataformas de interoperabilidad utilizan los Web Services como las piezas de software de integración con las entidades gubernamentales y las empresas.

3.3. Estándares soportados

Los Web Services ofrecen comunicación para sistemas con diferentes sistemas operativos y plataformas de desarrollo. Para conseguir este objetivo, se basan en una familia de especificaciones de protocolos de industria, un conjunto extendido de estos estándares se denomina WS-*. Para la versión 1.0 de la Plataforma de interoperabilidad los estándares de relacionados con Web Services soportados son (IS21, 2011):

- XML, Version 1.1
- XML Schema, Versión 1.1
- SOAP, Versión 1.1
- WSDL, Versión 1.1
- XSL, Version 1.0
- WS-Security 1.0

Con el cumplimiento de los anteriores estándares y versiones, se dice que el Web Service cumple con la especificación Basic Profile 1.0 de WS-I.

3.4. Creación del trámite o servicio

La siguiente gráfica muestra la secuencia ocurre en el proceso de registro del trámite o servicio en la Plataforma de interoperabilidad, así como el responsable de cada tarea en el flujo como se muestra en la Figura 2. Diagrama de Secuencia de un sistema de trámites por actores.

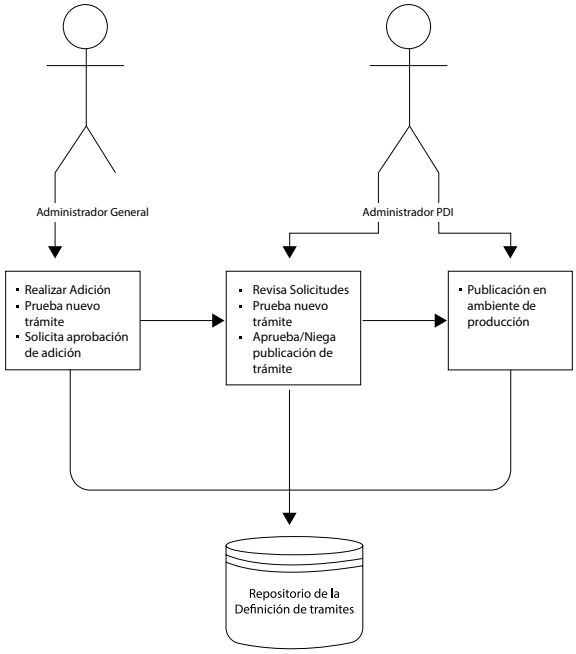


Figura 2. *Diagrama de Secuencia de un sistema de trámites por actores.*
Fuente: *Elaboración propia*

Arquitectura tramitador en línea

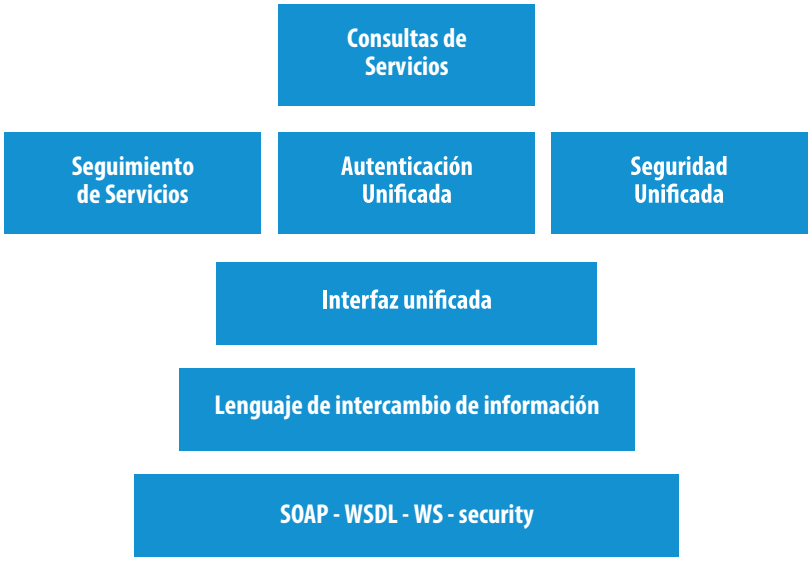


Figura 3. *Arquitectura Tramitador en Línea*
Fuente: *Elaboración propia*

La figura 3 muestra los componentes de la arquitectura del Tramitador en Línea, funcionalmente la herramienta debe permitir a los ciudadanos crear un usuario y darse de alta en el sitio, ya dentro el ciudadano contará con una serie de servicios disponibles según su perfil y ciudad seleccionada, cada servicio o trámite puede iniciarse desde la plataforma, entregando a la ciudadanía la oportunidad de gestionar sus trámites con el estado sin necesidad de someterse a desplazamientos a una ventanilla presencial e incurrir en gastos de transporte (Solano et al, 2017). En este orden de ideas las entidades del orden territorial deberían estar diseñando su catálogo de servicios y trámites sobre la plataforma mencionada, resultado que no ha sido el esperado, dado que para el diseño de un trámite es necesario asumir una curva de aprendizaje demasiada larga y conseguir recurso humano demasiado especializado (Figueroa, Porto, Simancas & Higuera, 2017).

La explosión que ha surgido en torno a la demanda se servicios web, ha llevado a construir plataformas muy flexibles diseñadas para cambiar, las entidades del sector público son entidades donde el cambio es muy complicado el MinTic previendo que Tramitador en Línea podía quedarse desactualizado en cuestión de meses, diseñó un lenguaje de intercambio de información, el cual a primera vista resulta un aporte interesante e intenta resolver la sostenibilidad de la herramienta en el tiempo.

El Lenguaje para el intercambio de información (LEI) utilizado por Bertram y Kleiner (2012), es un estándar que usa el Gobierno en línea basado en XML, que permite que los distintos sistemas de información puedan sintonizar y permutar información de forma apropiada, sin tener que renunciar a los sistemas actuales.

En síntesis, este lenguaje intenta describir todos los posibles objetos existentes en los sistemas de información ejemplo:

Vehículo		**Ciudadano**	
Placa	Cadena	Identificación	Cadena
NoChasis	Cadena	Nombre	Cadena
NoMotor	Cadena	Apellido	Cadena
...	...	...	...
Color	Cadena	Edad	numérico

Figura 4. *Diccionario de datos agrupado por etiquetas XML*
Fuente: *Elaboración propia*

De esta manera no importa el lenguaje los objetos están definidos en un diccionario de datos y agrupados en etiquetas XML, asegurando la interoperabilidad de la información. La gran dificultad de esta solución se presenta en la ambigüedad que pueden presentar algunos objetos para diferentes entidades del orden territorial, por ejemplo, objetos que definan Impuestos territoriales, estos son muy diferentes para cada entidad territorial, y uno de los usos con mayor potencial de crecimiento es la de pagos de impuestos en línea. Podríamos construir una lista de inconvenientes a la implementación de la solución, solo es de resaltar el crecimiento de la demanda de servicios web desde dispositivos móviles y estos prefieren servicios web tipo RESTful, tipo de servicios no disponibles en la plataforma.

El tramitador en línea propuesto por el MinTic no responde a la necesidad de exponer el catálogo de trámites y servicios a los ciudadanos.

4. Elementos de integración para un web service

Para el desarrollo de una web Service para la implementación de un tramitador de servicios en línea se deben tener los siguientes elementos:

- Proyecto Eclipse Plantilla para la creación de WebService con un único método, donde el formato de entrada y salida depende de los XSD definidos por el desarrollador del Web Service o utilizando como base el repositorio de XSD del lenguaje de intercambio de información GEL-XML
- Interfaz plantilla a implementar para crear Acciones para el WebService de forma separada.
- Clases Bindings generadas a partir del XSD de entrada y salida definido por el desarrollado por cada operación (la generación se puede hacer a través de Netbeans o Eclipse)
- Procedimientos y Documentación para la utilización del Proyecto Eclipse Plantilla, la Interfaz ExecutableAction y la generación de clases Java a partir del XSD.

Desarrollador del Proceso:

- Proyecto Eclipse Plantilla para la creación de Proxies que consuman los WebServices expuestos, y a ser consumidos por el proceso de Metalogo.
- WSActionHandler empaquetado como librería JAR utilizado para la comunicación desde el proceso hacia el Proxy.
- Clases Bindings generadas a partir del XSD de entrada y salida definido

por el desarrollado por cada operación (la generación se puede hacer a través de Netbeans o Eclipse).

- Nota: estas clases pueden ser compartidas por el desarrollador del WebService o pueden ser generadas por el mismo desarrollador del Proceso.
- Procedimientos y Documentación para la utilización del Proyecto Eclipse Plantilla para la creación de Proxies, la librería WSActionHandler y la generación de clases Java a partir del XSD.
- Procedimientos y Códigos Base del fragmento de proceso desde el cual se hace el llamado al WebService:
    1. Códigos Java base para establecer las variables de petición para el llamado del WebService.
    2. Definición de parámetros para la invocation del WSActionHandler dentro de un elemento nodo en el proceso.
    3. Códigos Java Base para obtener las variables de respuesta retornados por el llamado al WebService.

Un último elemento incorporado a la solución es el de la seguridad, para esto se utiliza WS-Security, particularmente las especificaciones XML Encryption, XML-Signature y WS-UsernameToken.

5. Conclusión

Es importante invertir en equipos de investigación y desarrollo, apostarle a construir nuevos conocimientos es siempre ganador, y es de admirar el esfuerzo realizado por las diferentes organizaciones en busca de proponer un portafolio de trámites y servicios en línea, pero es necesario anotar algunas recomendaciones para el proceso de construcción, puesta a punto y mantenimiento.

No enfocar los esfuerzos del equipo en construir nuevos estándares, esto puede traer un gran desgaste y el invitar a los grandes de la industria a utilizar los mismos puede ser bastante tedioso.

El partir de un motor que cuente con el soporte de una comunidad reconocida que se responsabilice por la mejora continua suma puntos en la elección.

Usar granularidad fina en el diseñar la solución, esto permite cambiar componentes mitigando el impacto que puede generar la acción en mención.

Es una necesidad la implementación de portafolio de trámites y servicios en línea más cuando pensamos en expandir las fronteras de la organización que busque su montaje.

Referencias Bibliográficas

Bertram, J. & Kleiner, C. (2012). Secure Web Service Clients on Mobile Devices. Procedia Computer Science, 10, 696-704. https://doi.org/10.1016/j.procs.2012.06.089

Figueroa, M., Porto, R., Simancas, R. & Higuera, V. (2017). Modeling demand for air cargo in the Colombian context FULL.doc. En ResearchGate-Science and Civil Engineering (AMSCE 2017) (p. 6). https://doi.org/10.2991/amsce-17.2017.31

Gobierno Ciudad de México. (2017). Portal de Trámites CDMX. Recuperado 14 de noviembre de 2017, a partir de: http://www.tramites.cdmx.gob.mx/

IS21. (2011). Marco para la Interoperabilidad del Gobierno en línea Estrategia de Gobierno en línea de la República de Colombia. Recuperado a partir de: http://www.programa.gobiernoenlinea.gov.co/apc-aa-files/5854534aee4eee4102f0bd5ca294791f/CAPACITACION_ENTIDADES_ENRUTADOR.pdf

Kennedy, S., Stewart, R., Jacob, P. & Molloy, O. (2011). StoRHm: a protocol adapter for mapping SOAP based Web Services to RESTful HTTP format. Electronic Commerce Research, 11(3), 245-269. https://doi.org/10.1007/s10660-011-9075-3

Lessen, T., Nitzsche, J. & Leymann, F. (2009). Conversational Web Services: leveraging BPELlight for expressing WSDL 2.0 message exchange patterns. Enterprise Information Systems. https://doi.org/10.1080/17517570903046300

Mohammadi, F. A., Nematbakhsh, N. & Nematbakhsh, M. A. (2012). WSDATool: A Novel Web Service Developer Assistance Tool Using a New Complementary Service Publishing Method. Modern Applied Science, 6(9), 42. https://doi.org/10.5539/mas.v6n9p42

Pautasso, C. (2009). RESTful Web service composition with BPEL for REST. Data & Knowledge Engineering, 68(9), 851-866. https://doi.org/10.1016/j.datak.2009.02.016

Pinzon, C., Bajo, J., De Paz, J. & Corchado, J. (2011). S-MAS: An adaptive hierarchical distributed multi-agent architecture for blocking malicious SOAP messages within Web Services environments. Expert Systems with Applications, 38(5), 5486-5499. https://doi.org/10.1016/j.eswa.2010.10.088.

Porto, R., Sarmiento, A., Garcés, L. H., Solano, A. F. P., López, D. & López, A. (2017). Marco de enseñanza basado en ovas para el adiestramiento en la introducción de los algoritmos (olia). Etic@ net, 1(17), 60-75.

Porto, R. S. (2016). Software móvil para la enseñanza de las Normas Internacionales de Información Financiera (NIC-NIIF) APPIFRS (parte 2). Journal of Engineering & Technology (2256-3903), 5(1).

Solano, R. P., Barceló, R. P., Gómez, A. C., Barbosa, J. C., Gutiérrez, C. E., De los Ríos, J. & Meza, E. H. (2017). Framework ágil para el control de recetas médicas que utiliza la tecnología NFC (FARM). Revista Lasallista de Investigación, 14(1).

APLICACIÓN DE UNA ESTRUCTURA CONCEPTUAL PARA EL DESARROLLO DE SISTEMAS DE INFORMACIÓN EN EL CONTEXTO DE CIUDADES INTELIGENTES•

Danilo García Salas *
Wilson Nieto Bernal**
Roberto Porto Solano ***
Martin Diaz Mora****

• Capítulo de libro resultado de investigación. Producto Colaborativo de los grupos de investigación Aglaia y GReCIS de la Universidad del Norte. Producto resultado del proyecto institucional de la universidad del Norte junto con Instituto de Estudios Hidráulicos y Ambientales – IDEHA en colaboración con la Corporación Universitaria Americana (2015).

* Ingeniero de Sistemas. Magister en ingeniería de Sistemas y Computación. Corporación Universitaria Americana. Correo: garciadanilo@coruniamericana.edu.co
** Ingeniero de Sistemas. Magister en Ciencias de la Computación. Doctor en informática y tecnologías de la información. Universidad del Norte. Correo: wnieto.@uninorte.edu.co
*** Ingeniero de Sistemas. Magister en ingeniería de Sistemas y Computación. Corporación Universitaria Americana. Correo: rporto@coruniamericana.edu.co
**** Ingeniero de Sistemas. Magister en ingeniería de Sistemas y Computación. Corporación Universitaria Americana. Correo: mdiaz@coruniamericana.edu.co

Resumen

Este documento presenta un marco establecido en 6 capas, bajo el Contexto de ciudades emergentes, para establecer las estructuras de información que los usuarios proporcionan a las personas, notificaciones de incidentes en tiempo real mediante los dispositivos IoT, quienes se encuentran estratégicamente ubicados para el almacenamiento periódico de información. El marco fue validado con la implementación de una red pluviométrica en la ciudad de Barranquilla y la red de aguas pluviales para la notificación a los ciudadanos a través del móvil. Terminales de los estados lluviosos de la ciudad, así como información histórica sobre las lluvias.

Palabras clave: Palabras clave: Framework, ciudades emergentes, dispositivos ubicuos, nube informática.

Abstract

This document presents a framework established in 6 layers, under the context of emerging cities, to establish the structures of information that users provide to people notifications of real-time incidents by means of the devices IoT, who are strategically located for the periodic storage of information. The framework was validated with the implementation of a pluviometric network in the city of Barranquilla and the rainwater network for notification to citizens through mobile terminals of the city's rainy states, as well as historical information about of the rains.

Keywords: Framework, Emerging Cities, Ubiquitous Devices, Cloud Computing.

Introducción

Actualmente muchos han sido los aportes que se han realizado en los sistemas de información con sensores ubicua, estos toman vital importancia debido a la existencia de la tendencia "IoT (Internet of Things, Internet de las cosas)", la cual ha traído consigo que se abran nuevas propuestas relacionadas a la forma como se diseñarán y se desarrollan propiamente sistemas de información, que permitan monitorear este tipos de arquitecturas, las que poseen un conjunto de dispositivos para realizar mediciones a través de sensores y las ubicaciones en donde se encuentra estratégicamente localizados.

Diferentes factores como el volumen y la velocidad de transferencia de los datos que se generan en este tipo de redes son un aspecto relevante, por lo que se deben tener en cuenta excelentes prácticas, así como diseñar las estrategias más pertinentes para mantener las consistencias de estos y en caso de requerirse, entregar los datos en tiempo real.

Otros de los aspectos que está tomando mucha importancia recientemente y complementado con el internet de las cosas trae consiga una amplia y más variada oferta de servicios, esta tendencia es la de la computación en la nube "Cloud Computing". Al igual muchas de las aplicaciones que se les dan a todas estas tendencias se aplican en las ciudades emergentes, para ofrecerles a sus ciudadanos una mejor calidad de vida, con todos los beneficios que se le pueden entregar con la aplicabilidad de todas las tecnologías en mención. Al igual uno de los aspectos por los cuales se justifica la aplicabilidad de tecnología en las ciudades emergentes, es la óptima utilización de recursos no renovables, los cuales, en su escasez y costos de administración dentro de las ciudades, justifican que se le aplique tecnología para tomar las decisiones más pertinentes al respecto, así mismo como optimizar los costos y realizar poco derroche de estos.

2. Materiales y métodos

La implementación de sistemas de información que de alguna u otra forma ayudan a capturar variables de áreas como la agricultura, la arquitectura (gestión de edificios), salud, transporte y energía, que son de vital importancia para la toma de decisiones tanto a presente como a futuro.

En Australia y en el Reino Unido, realizaron un estudio el cual muestra cómo los ciudadanos participan en proyectos relacionados con proyectos mediados por las Tic's, en tanto que aclaran que actualmente nos encontramos en la era de la computación ubicua, ya que casi que cualquier ciudadano cuenta con un dispositivo móvil, el cual generalmente está alimentando de información a sistemas más robustos o centralizados, y los volúmenes son bastante grandes, al igual sugieren una taxonomía de aspectos que se deben tener en cuenta para que los ciudadanos participen en proyectos urbanos mediados por TIC's (Luo, Fang & Huang, 2015).

Por otra parte en Londres y Australia, realizaron unos estudios con ciudadanos, mediante los cuales a través de aplicaciones instaladas en dispositivos móviles, podían participar en actividades de participación ciudadana, lo cual permitió crear una taxonomía, en la cual variables como el impacto que tendrían sobre los ciudadanos son tomadas en cuenta, y lo cual fue posible debido a la gran cantidad de aplicaciones y el uso que le fue realizado a los ciudadanos (Salim & Haque, 2015). Este estudio es una fiel muestra de que a través de la tecnología se pueden generar y promover aspectos culturales, sobre los habitantes de una ciudad (Sepúlveda et al. 2018).

La gran creciente tendencia del internet de las cosas, ha traído consigo un gran sinnúmero de aplicaciones, primordialmente teniendo inconvenientes de interoperabilidad entre varias entidades del IoT, lo cual hizo posible la creación de una arquitectura que tuviera en cuenta este aspecto más: automatización, inteligencia y cero configuraciones, son algunos de los aspectos más importantes de esta arquitectura, cuyo nombre es DIAT (Distributed Internet-like Architecture for Things), con lo que previamente se hizo una revisión del estado del arte y más aún, revisión de lo que utilizan los proyectos más notables concernientes al IoT y arrojando una arquitectura con las características antes mencionados (Sarkar, Nambi, Prasad & Rahim, 2014; Sarkar, Prasad, Rahim, Neisse & Baldini, 2015).

Uno de los aspectos que ha llamado notablemente la atención al internet de las cosas, es sí las actuales infraestructuras que poseen los proveedores de Internet son suficientes para soportar esta tendencia que crece exponencialmente, por tanto este estudio (Chang, Chen, Chen & Chao, 2012), muestra un análisis en el que demuestra que la actual tecnología no es suficiente para que mantengan calidad en las conexiones, por lo cual hacen una propuesta preliminar y la cual muy necesaria, para el prospecto esperado de conexiones, en un futuro muy cercano.

En Italia, propusieron un marco de trabajo integrado para configurar dispositivos inalámbricos a través de servicios web, lo cual lo lograron usando interfaces compartidas entre los dispositivos y la arquitectura que se desplegó en la mesa de pruebas tomada para validar el trabajo en mención. Al igual este marco de trabajo se validó con las guías de SENSEI, el cual es el mayor referente de los sensores de redes inalámbricas, lográndose uno de los mayores avances en interoperar con futuras tecnologías que se propongan, como uno de los aspectos más importantes de este marco de trabajo (Castellani et al. 2010). Esta misma propuesta se ha extendido hasta tal punto que su aplicabilidad se le ha dado a aspectos urbanos, en la que aplicabilidad del IoT a ciudades es más pertinente para mejorar la toma de decisiones, realizan una análisis de tecnologías y estándares, así como una breve aplicabilidad de las mediciones que se pueden realizar con una red inalámbrica de sensores (Zanella, Bui, Castellani, Vangelista & Zorzi, 2014). Al igual este otro trabajo muestra la ciudad de Padova, Italia, como una ciudad piloto para realizar pruebas de IoT, en la índole urbana (Cenedese, Zanella, Vangelista & Zorzi, 2014).

Uno de los aspectos que al igual se está desarrollando amplios estudios, al igual que una tendencia derivada del internet de las cosas IoV "Internet of Vehicle", es llevar conectividad de redes a vehículos móviles, por tanto, en (Salahuddin, Fuqaha & Guizani, 2015), realizan una propuesta para comunicar vehículos, al igual que aplicando aspectos de optimización, es específico de heurísticas, ya que el problema es de índole multiobjetivo, para realizar las correspondiente reconfiguraciones, de tal forma que se puedan desplegar los servicios a los clientes, sin importar de cierta manera la demanda, ya que la reconfiguración es consecuente con ésta.

Existen otras propuestas como Abdelwahab, Hamdaoui, Guizani y Rayes (2014), proponen CARS: Cloud-assisted remote sensing, sobre el que se usan las ventajas de la computación en la nube, para sensar datos de distintos emisores, sobre los cuales se puedan obtener, procesar y persistir la información, con las tecnologías de la nube, entre las características más relevantes encontramos: elasticidad, pago sobre demanda, interoperabilidad tecnológica, entre otras características.

Por otra parte, análisis de la perspectiva del IoT desde China, se han realizado, notándose un aspecto muy importante, en tanto que muchas han sido las ayuda que el Gobierno Chino, ha ofrecido para fortalecer esta tendencia, ya que son muchas las ventajas que ofrecen, al igual que destacan dos aspectos sobre los cuales han mostrado trabajos: minería y

transporte inteligente tal como ellos lo han notado (Chen, Xu, Liu, Hu & Wang, 2014).

Esta otra propuesta muestra cómo Kantarci y Mouftah (2014), una arquitectura desplegada puede ayudar con la seguridad pública, en donde los sensores, están conectados a dispositivos móviles, y en la que una arquitectura como la que proponen se puede utilizar para prevenir y controlar desastres.

Como bien se sabe, el esquema convencional del Internet, no fue diseñado para soportar dispositivos ubicuos, que es primordialmente sobre donde se enmarca el IoT, por tanto esta propuesta (Yue, Guo, Li, Asaeda & Fang, 2014), es una nueva propuesta de arquitectura la cual llaman Data Clouds, en la que se comparte mucha información sobre diferentes aspectos, y que se enmarca bajo el enfoque de las ciudades inteligentes, en las que existe una amplia y variada gama de servicios enfocados para los ciudadanos, y entidades que necesitan información de muchos de los datos que se recogen a través de cada uno de los dispositivos, que se tienen configurados para realizar el sensado.

Una de las aplicaciones que son bastantes importantes en el internet de las cosas (IoT), es la aplicabilidad en el sector industrial, como en (Qiu, Luo, Xu, Zhong & Huang, 2015), en donde específico es utilizado para la administración de activos en el sector logístico, articulando se con las empresas manufactureras sobre un parque industrial. Este otro trabajo (Guo, Ngai, Yang & Liang, 2015) ayuda en el estado de la producción, dentro de una planta industrial, en el que dependiendo del estado en donde se encuentra la producción, ayuda a reorganizar el trabajo y mejorar los índices de desempeño y optimización de los recursos, tanto de máquinas como en el Personal que labora, permitiendo tomar mejores decisiones.

En este otro trabajo Calvillo, Sánchez-Miralles, y Villar en 2016, dan los lineamientos más importantes que se deben tener en una ciudad inteligente para que se planteen los modelos más pertinentes para la administración y planificación de energía, ya que siendo éste un recurso muy vital para una ciudad, y más que contextualizando sobre ciudades inteligentes es pertinente tomar en consideración este tipo de modelos.

Las soluciones tecnológicas que capturan procesan y muestran pronósticos que ayudan a la toma de decisiones en el contexto de ciudades

emergentes como el presentado por García, Porto, Simancas, Molina y Hernández (2016).

Por otro lado, los Sistemas de información permiten mejorar los puntos de atención de las instituciones prestadoras de servicios de salud, reducen drásticamente los errores de prescripciones como el presentado en (Solano et al, 2017).

Este otro trabajo Chin, Kim, Heo, y Jang (2015), realizan una propuesta a futuro de cómo se deberían desplegar, servicios a través del Internet de las cosas IoT, al igual que la infraestructura que debe ser tomada en cuenta, para que exista la mayor optimización en el despliegue de los servicios.

Sabiendo que el agua potable es recurso vital para las sociedades, se diseñó una propuesta a través de la cual usando dispositivos del internet de las cosas se recolectaron datos a través de toda la red, para de esta manera tomar decisiones sobre posibles fugas o pérdidas del recurso, y tomar las indicaciones o decisiones pertinentes (Koo, Piratla & Matthews, 2015).

Esta otra propuesta aplica para dos aspectos fundamentales del internet de las cosas: una es la recolección de los datos y la otra es como esos datos se tratarán para aplicarle a estos minados de datos y en base a esto poder tomar las mejores decisiones, y sobre todo con las tendencias escalables que son propias de la computación en la nube (Villalba, Pérez, Carrera, Pedrinaci & Panziera, 2015).

3. Marco de trabajo fracise

Uno de los aspectos más importantes que cabe destacar de este framework es que realizado una revisión bibliográfica en Colombia, es el primer Framework de este tipo que se proponen para los propósitos de ciudades emergentes y ciudades inteligentes, el cual es un referente muy importante para muchas de nuestras ciudades de nuestro país Colombia. Por mencionar una de las problemáticas que afronta la ciudad de Barranquilla, la cual son los arroyos, cuyas pérdidas materiales son bastante considerables, así como en muchos casos ocurren pérdidas de vidas, lo cual es muy lamentable. He aquí la importancia de trabajar en este tipo de trabajos lo cual repercute enormemente en el mejoramiento de la calidad de vida de los ciudadanos de alguna manera.

El framework propone 6 capas, las cuales son fundamentales para desarrollar un sistema de información en el contexto de las ciudades emergente. Hay un aspecto en particular que es transversal a todas las capas que propone FRASICE, y es la seguridad y la capa de rendimiento, que es transversal en 3 capas del framework vistas en la Fig. 1. A continuación describiremos cada uno de los aspectos fundamentales del framework, así como cada uno de los detalles y características de cada una de sus capas:

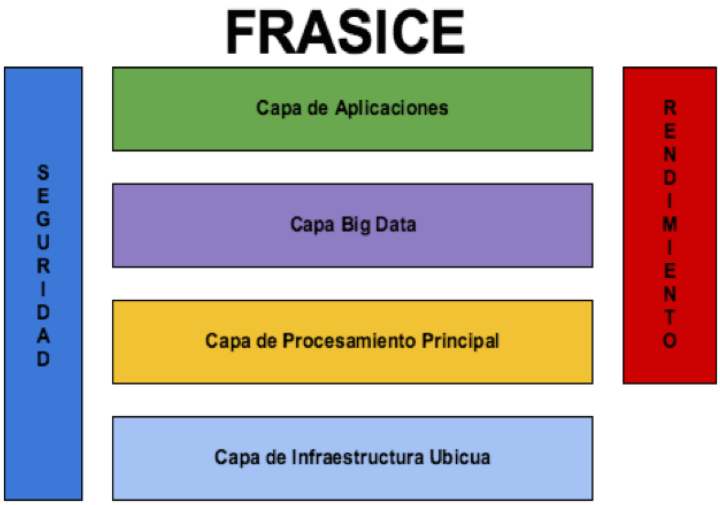


Figura 1. *Capas del marco de trabajo FRASICE.*
Fuente: *Elaboración propia*

3.1. Capa de Seguridad

Como bien antes fue mencionado, es quizás, uno de los aspectos más importantes, ya que de una u otra forma es necesario que los datos estén seguros y se entreguen a los usuarios, sin ningún tipo de inconvenientes. Al igual cabe resaltar que cualquier aspecto que sea necesario denotar sobre cada capa referente a la seguridad, se hará con detalles cuando sea pertinente. Por lo pronto, el referente al que hace referencia la seguridad es en el aspecto global, por decirlo de esta manera.

3.2. Capa de Infraestructura Ubicua

Esta capa hace referencia al conjunto de instrumentos, mediante los cuales se realizan las mediciones, los sensores. Generalmente se localizan en sitios estratégicos, en algunos casos, en lugares donde el foco de atención es mayor o en mejor de los casos, donde las variables de sensado que

posean, sean las variables que se deseen medir en ese punto en específico. Preferiblemente sobre un punto se colocarán varios sensores los cuales, con un tramado generalizado permitirá comunicarse correctamente con el sistema de información, ya que entre estos se encuentran sincronizados, por marcas, id's los cuales permiten conocer la procedencia de los datos, a nivel lógico por decirlo de esta manera es necesario, que cada estación posea un identificador unívoco, así como cada uno de los sensores que posee adjunto a la estación, y en la cual llevarán cada vez que se comunique con el sistema de información, al lado de los datos medidos por la estación. Preferiblemente es mejor que los dispositivos hardware, sensores, se estén comunicando con una microcomputadora, la cual a su vez posea un software, el cual permite las configuraciones básicas de las estaciones, como lo son, por mencionar algunas: Dirección Ip y/o dominio end-point del sistema de información, periodicidad de toma de datos para cada uno de los sensores, tipo de envío hacia el end point (UDP/TCP), configuraciones de id's (identificador unívoco estación y de los sensores, sincronizados con el sistema de información), entre otras configuraciones que son importantes. Preferiblemente es mejor tener un aplicativo web embebido sobre cada una de las estaciones, para realizar todas estas configuraciones, de lo contrario un dispositivo electrónico, acompañado de micro-controladores PIC's que reciban instrucciones mediante una conexión TCP/UDP, y se le envíe un tramado que permitan hacer cierta configuración sobre la estación, y sea esta quien interprete la instrucción o las instrucciones. Para comunicarse desde las estaciones hacia los sistemas de información el canal puede hacerse mediante internet móvil, si existe cobertura de alguna red celular, en caso contrario, es una red MAN, con infraestructura inalámbrica multipunto.

La energía con que se alimentarán estas estaciones, son con energías de tipo renovables, paneles solares que alimentarán a una batería, preferiblemente, si poseen energía de la red eléctrica, también es posible hacerlo.

Seguridad. Con lo referente a la seguridad en esta capa viene dado a que preferiblemente, es mejor que las estaciones se conecten al sistema de información mediante un túnel VPN, en caso de enviar los datos por Internet. Si la conectividad se realiza mediante dispositivos MAN punto a punto, comunicándose desde las estaciones hacia el sistema de información, configurar los equipos, con la seguridad pertinente para que no se fugue información alguna.

3.3. Capa de Procesamiento Principal

La capa de procesamiento principal hace referencia a aquella infraestructura (servidor(es)), mediante los cuales se recibe información desde la capa de infraestructura ubicua, se persiste en la capa de Big Data y se sirve o mejor se presentan los datos para la capa de aplicaciones. Por lo tanto, la importancia o el rol que posee esta capa es bastante importante, en tanto que en lo mejor posible es mejor independizar totalmente esta capa, de la capa de Big Data, debido a que esta capa se puede volver tan elásticamente posible dependiendo de la demanda de usuarios haciendo solicitudes al sistema de información, este se logra colocando balanceadores de carga sobre el clúster de servidores disponibles.

Es pertinente colocar esta capa sobre infraestructuras del cloud computing, la cual se pueda colocar en condiciones de escalabilidad automática, por lo cual la experiencia de usuario no se verá afectada por ningún motivo, ya que todo el requerimiento de infraestructura se pueda suplir dependiendo de la cantidad de usuarios que requieran la aplicación, lo cual antes era impensable, ya que adaptar una infraestructura in-situ no es muy fácil como hacerlo en tecnologías basadas en el cloud computing.

Otros de los aspectos que es bastante importante es que maneje servicio para comunicación síncrona con los clientes, bien sean dispositivos móviles o de escritorios a través de un navegador, en caso de que sea necesario, dar una orden en tiempo real, que pueda realizarse, sin ningún tipo de inconveniente.

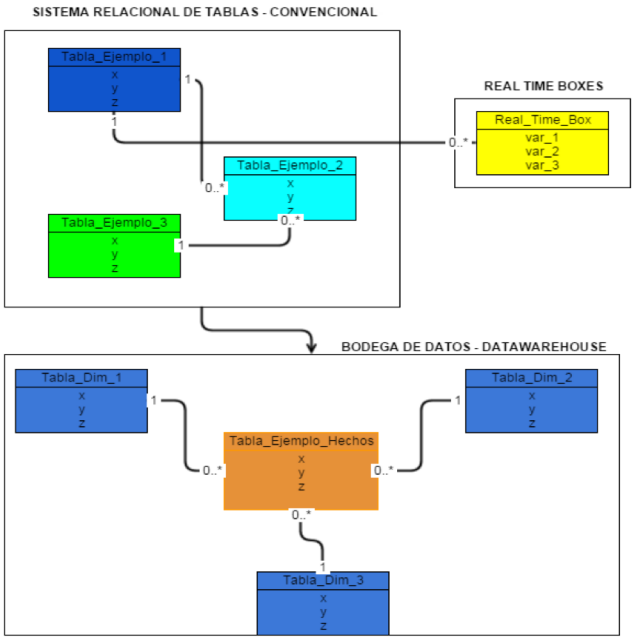


Figura 2. *Real Time Boxes Ejemplo de los Real Time Boxes, "Cajas de Tiempo Real", las variables que contienen estas tablas son variables que son de alto interés para los usuarios que usan el sistema de información.*
Fuente: *Elaboración propia.*

Seguridad. Entre los aspectos que es necesario tener en cuenta para la seguridad en la capa principal de procesamiento, en el caso de utilizarse tecnologías en el cloud, los accesos a las consolas de administración deben configurarse con el máximo nivel de seguridad, a los usuarios pertinentes delegar las permisividades que sean necesarios para los mismos y al igual tener en cuenta no sólo proteger los accesos a consola con usuario o contraseña, además colocarle otro tipo de autenticación.

Rendimiento. El rendimiento de esta capa se puede medir por el porcentaje de utilización de la máquina(s), que hagan parte de esta. Al igual en caso de utilizarse infraestructura de la nube, se podría pensar en una política de elasticidad.

3.4. Capa Big Data

Esta capa es la concerniente para persistir los datos en bases de datos, no se casa en una tecnología en específico, bien pueden ser bases de datos relacionales o bien no relacionales. Aunque muy distinta a la capa de procesamiento principal, esta capa debe garantizar continuidad y

disponibilidad, lo cual se logra si el servicio de bases de datos del que se disponga posea esta capacidad, al igual como en la capa de procesamiento principal es necesario que se pueda modificar esta capa, según la demanda requerida de información. Otro de los aspectos importantes es como existen variables acumuladas que se mostrarán con mayor frecuencia, dado la alta granularidad de los datos, se hace pertinente, crear un mecanismo para obtener esas variables de forma más rápida, para lograrlo llamamos algo que nombramos Real Time Boxes (Cajas de Tiempo Real), no es más que una tabla que guarda los valores acumulados para ser mostrados a la capa de aplicaciones, la cual no requiere de robustas consultas que degradan el desempeño del sistema de información, mejor aún se puede pensar en colocar los real time boxes, en bases de datos alojadas en memoria RAM, lo cual es un servicio del que disponen los servicios del cloud. Por otra parte, la bodega de datos es otro aspecto fundamental, en este tipo de diseños, por la alta granularidad de los datos, es muy común que, por parte de los administradores del sistema, se obtengan una cantidad de reportes, los cuales se obtendrían propiamente de la bodega de datos como se muestra en la Fig. 2.

Seguridad. Es preferible que esta capa se encuentre en un área que no se acceda desde el mundo, por todo lo contrario, las peticiones de sentencias DML, sólo deben configurarse desde las capas de procesamiento principal y desde la capa de aplicaciones únicamente.

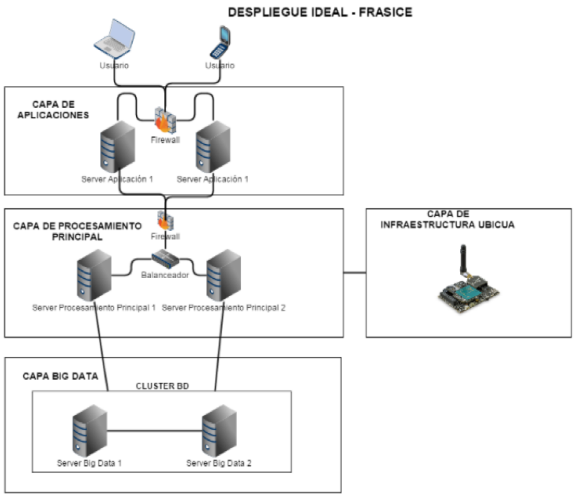


Figura 3. *Despliegue Ideal de las capas de FRASICE. Las capas de Seguridad y de Rendimiento se encuentran inmersas en las capas de Aplicaciones, Procesamiento y Big Data.*
Fuente: *Elaboración propia.*

Rendimiento. Aunque modificar esta capa no es muy necesario, debido a que cuando se coloca se fijan los recursos en la capa de infraestructura ubicua, se tienen un estimado del volumen de peticiones hacia el servidor. En algunos casos es pertinente realizar políticas de escalabilidad, si bien se tiene esta capa sobre tecnologías en la nube.

Peticiones hacia el servidor, ya que en algunos casos es pertinente realizar políticas de escalabilidad si bien se tiene esta capa sobre tecnologías en la nube.

3.5. Capa de Aplicaciones

En la capa de aplicaciones se tienen en cuenta todas las aplicaciones que lo usuarios necesitan utilizar, desde la de usuarios finales, así como las de usuarios de tipo administrador, o aquellas aplicaciones para administrar las bodegas de datos, en este caso, la conexión es pertinente realizarla hacía la bodega de datos. Para el caso de las aplicaciones de usuario final, se conectan directamente con la capa de procesamiento principal, ya que en lo posible la capa de procesamiento principal, entrega los datos en crudo (.xml y/o. json), para luego el correspondiente formato sea realizado desde las aplicaciones.

Se recomienda, para las aplicaciones que se desarrollen a la medida, llevarlas a cabo bajo el enfoque MVC (Modelo, Vista, Controlador), y de cierta manera permitan disminuir la capacidad de procesamiento, por parte de la capa de procesamiento principal y más bien se comparta el procesamiento con el cliente, y baje la carga en los servidores como se muestra en la figura 3.

Seguridad. Como es pertinente utilizar una infraestructura de la nube, de tipo elástica, es más que necesario tomar las precauciones que antes mencionábamos para la capa de procesamiento principal.

Rendimiento. Para el rendimiento es igual de necesario tomar las indicaciones dadas para la capa de procesamiento de datos principal.

4. Resultados

Teniendo en cuenta la aplicación del framework FRASICE, se enmarcó en el proyecto: Desarrollo de proyectos I+D en hidro sistemas estratégicos del Departamento del Atlántico y arquitectura de red para plataforma de información hidro ambiental del Departamento del Atlántico, como resultado

se desarrolló un sistema de información mediante el cual la ciudadanía pueda revisar en tiempo real, el estado de la red pluviométrica, la cual se encuentra instalada sobre la ciudad de Barranquilla y la cual avisa los lugares en donde se encuentra lloviendo.

Cada una de las capas que enuncia el framework FRASICE, se utilizaron para llevar a cabo este proyecto.

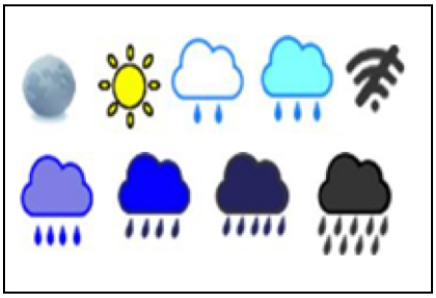


Figura 4. *Iconografía vista pública.*
Fuente: *Elaboración propia.*

4.1. Sistema de red pluviométrica

El sistema de información de la red de pluviometría se caracteriza por estar dividido en dos capas: Front End y Back End. La capa Front End hace referencia, a todo lo concerniente a interfaces gráficas e interacción con el usuario, y el back end, hace referencia a todo lo concerniente a los datos en crudos, que se obtienen de los sensores que se encuentran distribuidos a través de la red de pluviometría, estos se entregan mediante un servicio web bajo el formato y las tecnologías de Api restful, la cual a su vez utiliza la tecnología http para realizar peticiones al servidor (Inserción, actualización, eliminación, monitoreo de estaciones, reportes de administrador, listar sensores, listar usuarios, lectura). En el momento que la interfaz gráfica requiera una información de la red de pluviometría, realiza una petición hacía el servicio web, el cual entrega la información en formato. json, un estándar bastante reconocido para la portabilidad de información, por lo cual se optó como formato para transferir la misma.

En la vista pública, se definió la tabla de medida de precipitación e intensidad, con la finalidad de desarrollar la iconografía de la red de monitoreo y asociarla con la tabla, asociar los iconos a la tabla tiene como propósito que la información suministrada por la red de monitoreo sea entendible al usuario, que identifique el estado de un evento de lluvia como se muestra en la figura 4.

Figura 5. *Función Posición Actual y Tabla de medida e Iconografía*
Fuente: *Google Maps*

Teniendo en cuenta las necesidades del usuario se implementaron funciones como localizar la posición actual del usuario, esta es importante en caso de que el usuario desee saber si alrededor de su posición si se presenta un evento de lluvia y poder identificar la magnitud del evento. También fue implementada la función de alejar el mapa, centrando el mapa en barranquilla con el propósito de poder visualizar toda la red de pluviómetros y facilitar la navegación del usuario en la plataforma.

Además, se agregaron diferentes capas en el mapa, estas contienen diferentes vistas del mapa, se presentaron para definir la más conveniente para los usuarios.

Para el funcionamiento de esta vista pública se implementó la conexión del api con la vista, esta conexión nos permite enlazar la información que envían los pluviómetros con la vista para poder presentar la información en la red de monitoreo.

Para mejorar la simplicidad de la información presentada a los usuarios se decidió que la información a proyectar por estación es el nombre de la estación y la precipitación acumulada representada en estados, ya que con la iconografía de los eventos pueden ser comparados con la tabla de medidas, y visualizar la intensidad y precipitación, y además el icono de gráfica que nos enlaza a las gráficas de esta estación como se muestra en la figura 5.

Al seleccionar Reporte Gráficas, presentamos una gráfica anual visualizando la precipitación acumulada en cada estación mensualmente, además que listamos las estaciones para generar gráficas por estación como se muestra en la figura 6.

Codigo	Estación	Precipitación acumulada del día (mm)
BAJO	Barrio Abajo	NR
CHIQ	Chiquinquira	0.76
CIUD	Ciudadela	0.51
DELI	Delicias	0.00
PUEB	El Pueblito	0.76
CAMP	La Campiña	NR
CONC	La Concepción	0.00
VICT	La Victoria	NR
RIOM	Riomar	NR
SANT	Santa Helena	0.00
SOLE	Soledad	1.27
UNOR	Uninorte	NR

Figura 6. *Visualización reporte gráficas.*
Fuente: *Elaboración propia.*

En la opción de renuncia de responsabilidades, se informa y advierte que no somos responsables de que acciones tome el usuario con la información que proporcionamos, al igual que presentamos enlaces externos y no somos responsables de la información proporcionada en estos enlaces, esta vista la presentamos en una ventana emergente como se muestrea en la figura 7.

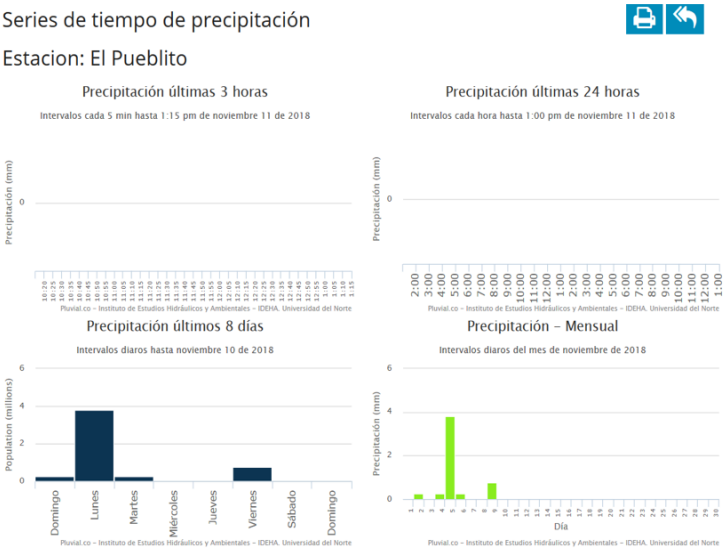


Figura 7. *Visualización de gráficas por estación*
Fuente: *Elaboración propia.*

5. Conclusiones

En cuanto al Framework plantea los aspectos más importantes que se deben tener en cuenta, a la hora de construir un sistema de información para ciudades emergentes, las cuales comprende la mayoría de las ciudades en vía de desarrollo que poseen la tendencia de brindar soluciones para sus ciudadanos. En cuanto a la aplicabilidad de esta infraestructura en un proyecto concreto, sería pertinente aplicar con la mayoría de los aspectos que se recomiendan en FRASICE, al igual en pluvial, se podría mejorar muchos aspectos, con harás de brindar mejores funcionalidades a la red pluviométrica.

Como bien es conocido, los sistemas de información actuales poseen fuertes tendencias hacía el Internet de las Cosas y con esto ocasionando una gran explosión de datos, por lo cual sería muy pertinente utilizar técnicas del Big Data, para poder procesar y almacenar la información que generan la gran cantidad de dispositivos que se colocan sobre las ciudades y algunos los cuales generan los ciudadanos, obteniendo datos diversos sobre los cuales se pueden realizar robustos análisis y realizar mejores tomas de decisiones.

Entre este mar de tecnologías y/o tendencias, pensar cuales serían los aspectos que se deben tener en cuenta para entregar sistemas de información robustos con altos estándares de calidad. Al igual la propuesta de las cajas de tiempo real como un mecanismo que aplica técnicas del Big Data, para entregar la información más importante en tiempo real, es un aporte muy valioso. Al igual que como todos los componentes que describe el Framework.

Referencias Bibliográficas

Amnistía Internacional Catalunya. (2009). Historia de los Derechos Humanos. 1ºEd. Barcelona: Grup d'Educación. Recuperado de: http://www.amnistiacatalunya.org/edu/pdf/historia/dudh-historia.pdf

Comisión de derechos humanos. (Diciembre 10 de 1948). Declaración Universal de los Derechos Humanos. Asamblea General de las Naciones Unidas. Resolución 217 A III.

Congreso de Colombia (26 de mayo de 1873), Código Civil. [Ley 57 de 1887].

Congreso de Colombia. (06 de enero de 2006). Ley por medio de la cual se modifican el código civil, la ley 84 de •1989, el código penal, el código de procedimiento penal y se dictan otras disposiciones. [Ley 1774 de 2006]. Recuperado de: http://es.presidencia.gov.co/normativa/normativa/LEY%201774%20DEL%206%20DE%20ENERO%20DE%202016.pdf

Congreso de Colombia. (12 de julio de 2011). Por la cual se dictan normas orientadas a fortalecer los mecanismos de prevención, investigación y sanción de actos de corrupción y la efectividad del control de la gestión pública. [Ley 1474 de 2011]. DO: 48.128

Congreso de Colombia. (24 de Junio de 2011), Artículo 34 [Capítulo I]. Por medio del cual se reforma el Código Penal, el Código de Procedimiento Penal, el Código de Infancia y Adolescencia, las reglas sobre extinción de dominio y se dictan otras disposiciones en materia de seguridad, [Ley 1453 de 2011]. DO: 48110.

Congreso de Colombia. (24 de Octubre de 2007). Ley de Responsabilidad Medioambiental. [Ley 26 de 2007]. Recuperado de: http://www.jausaslegal.com/resources/doc/071227-ley-resp-medioambiental-desembre07-5702.pdf

Constitución Política de Colombia [Const.] (1991), Art. 1 [Título I]. 2da Ed. Legis

Constitución Política de Colombia [Const.] (1991), Art. 79 [Título III]. 2da Ed. Legis

Constitución Política de Colombia [Const.] (1991), Art. 95 [Título V]. 2da Ed. Legis

Constitución Política de Colombia [Const.] (1991). 2da ed. Legis

Corte Constitucional Colombiana, Sala Plena. (18 de agosto de 2011). Sentencia C-620 de 2011. [M.P. JUAN CARLOS HENAO PÉREZ].

Corte Constitucional Colombiana, Sala Plena. (22 de junio de 2016). Sentencia C-327 de 2006. [M.P. GLORIA STELLA ORTIZ DELGADO].

Corte Constitucional Colombiana, Sala Plena. (Mayo 06 de 1998), Sentencia SU-182 de 1998. [MP Carlos Gaviria Díaz; José Gregorio Hernández Galindo].

Corte Constitucional Colombiana, Sala quinta de revisión. (Noviembre 20 de 2006), Sentencia T-959 de 2006. [MP Rodrigo Escobar Gil].

Corte IDH. (Febrero 5 de 2001). Sala Plena. Sentencia Caso "la última tentación de Cristo". Olmedo Bustos y otros vs. Chile.

Corte IDH. (Mayo 26 de 2010). Sala Plena. Caso Manuel Cepeda Vargas vs. Colombia.

Corte IDH. (Noviembre 23 de 2009). Sala Plena. Sentencia Caso Radilla Pacheco vs. México.

Corte IDH. (Septiembre 26 de 2006). Sala Plena. Sentencia Caso Almonacid y otros Arellano vs. Chile.

Fiscalía General de la Nación. (30 de diciembre de 2011). Resolución que crea y reglamenta la estructura y el funcionamiento de la Unidad Nacional de Fiscalías de Delitos contra los Recursos Naturales y el Medio Ambiente. [Resolución 3438 de 2011]. DO: 48.298.

Higuera, D, M. (2015). Protección de la Dignidad Humana: Control de Constitucional e implementación de Mecanismos Internacionales. Bogotá, Colombia: Grupo Editorial Ibáñez.

Kant, I. (1989). Metafísica de las costumbres. Segunda parte: Principios de la doctrina de la virtud. A. Cortina (trad.). Madrid: Tecnos.

Mesa, G. (2007), Derechos ambientales en perspectiva de integralidad: concepto y fundamento de nuevas demandas y resistencias actuales hacia el Estado ambiental de derecho. Bogota, Colombia.

Quinche, M. F. (2008). Derecho constitucional. Bogotá: Ibáñez.

Sabalain, C. (2009). Introducción de Conceptos Básicos, El Medio Ambiente y la Estadística. [Presentación].Recuperado de: https://www.cepal.org/deype/noticias/noticias/2/37052/2009_09_ma_id_37052_sabalian_cristina_ppt.pdf

Tobo, J. (2006). La Corte Constitucional y el control de constitucionalidad en Colombia. Bogotá: Ibáñez.

METODOLOGÍA PARA LA CREACIÓN DE ALGORITMOS EN LA EDUCACIÓN SUPERIOR•

Dougglas Hurtado Carmona *
Roberto Porto Solano**
Israel Escobar Hernandez***
Roberto Porto Barceló****
Ricardo Simancas Trujillo*****

• Capítulo de libro resultado de investigación. Producto Colaborativo de los grupos de investigación Aglaia y Gigetic del Politécnico Costa Atlántica. Producto resultado del proyecto institucional Metodología para la creación de algoritmos en la educación Superior (2017).

* Ingeniero de Sistemas. Magister en ingeniería de Sistemas y Computación. Doctor en Ciencias Mención Gerencia. Politécnico Costa Atlántico Correo: dhurtadoc@pca.edu.co

** Ingeniero de Sistemas. Magister en ingeniería de Sistemas y Computación. Doctor en ciencias. Corporación Universitaria Americana Correo: rporto@coruniamericana.edu.co

*** Ingeniero de Sistemas. Magister en gobierno de tecnología informática. Politécnico Costa Atlántica. dir_psistemas.edu.co

**** Ingeniero de Sistemas. Magister en ingeniería de Software y Sistemas informáticos. Politécnico Costa Atlántica. Correo: rportob@pca.edu.co

****** Licenciatura en educación. Magister en gerencia de mercadeo. Corporación Universitaria Americana Correo: rasimancas@coruniamericana.edu.co

Resumen

Todas las profesiones orientadas a las ingenierías tienen como requisitos en su proceso de formación cursar las materias de Algoritmia y Programación donde suelen ser vistas en primer y segundo semestre respectivamente. La mayor parte de los estudiantes presentan dificultades en el proceso de aprendizaje de la noción básica de los algoritmos. En este artículo se diserta sobre la propuesta metodológica para el desarrollo de algoritmos, tomando como base las bondades del diseño de algoritmos tradicional compuesto por 4 fases: Requisitos, Análisis, Diseño y Validación. Elementos propios del desarrollo basado en las distintas formas de programación como la Orientada a Objetos y patrones de arquitectura de software, son de suma importancia para el proceso enseñanza aprendizaje a partir del acrecentamiento de competencias en los algoritmos en la educación Superior.

Palabras clave: algoritmos, educación, aprendizaje, metodología, diseño

Abstract

All professions oriented to engineering have as requirements in their training process to take the courses of Algorithm and Programming where they are usually seen in first and second semester respectively. Most of the students in the learning process of the basic notion of algorithms. This article describes the methodological proposal for the development of algorithms, as a basis for the benefits of traditional algorithm design composed of 4 phases: requirements, analysis, design and validation. Elements of development based on different forms of programming such as Object Orientation and software architecture patterns, are of utmost importance for the teaching-learning process based on the increase of competences in algorithms in higher education.

Keywords: algorithms, education, learning, methodology, design

Introducción

En las instituciones de educación superior se viene presentando un fenómeno particular en la manera en que se desarrollan las competencias para el desarrollo de algoritmos, en especial en las careras que son soportadas por las ciencias de la computación, ingeniería del software y los sistemas de información. Este fenómeno consiste en que en que el diseño de algoritmos se expone como un proceso sin orden y sin control, validando la narración extravagante que al crear algoritmos se realiza de forma símil al edificar un edificio sin planificación al estilo ensayo y error.

En cierto sentido, se visualiza el diseño de algoritmos no se ejecuta como actividad ingenieril, donde se planifique y se haga gestión de los procesos para generar los entregables, más bien, se difunde como una labor creativa artesanal, sin una metodología clara, ni una forma de evaluar la funcionalidad, ni el cumplimiento en todos los escenarios del entorno en cuál se va a implementar (Alba, 2009). Lo que significa que la elaboración de algoritmos suele estar fundamentalmente en estado caótico (Arsham, 1995).

Las posibles razones por las cuales el diseño de algoritmos se encuentra en una crisis, se centra en la complejidad de las orientaciones de programación. Al aparecer nuevas formas de atracción de los datos, como la programación orientada a objetos (POO), ha generado una incertidumbre en la forma de enfrentar los problemas, en los cuales se debe hacer un diseño de algoritmos, dificultando la forma misma, y el costo de planificación, dificultando la s estimación en lo que se alude a recursos, en especial al tiempo (Aguilar, Muñoz & Poamres, 2010; Selic, Gullekson & Ward, 1994; Main & Savitch, 2001; Joyanes & Zahonero, 2002; Cohoon y Davidson, 2002).

Así pues, este estado en que se encuentra el desarrollo de algoritmos genera en la comunidad académica justificadas incertidumbres en relación con que, si las metodologías propuestas para el desarrollo son realmente ajustadas para esta intención, y también a la calidad de estos algoritmos que surgen, taxativamente se declara la intranquilidad de ¿qué nivel de calidad poseen?

La existencia de una crisis en el diseño de algoritmos se observa con mayor preocupación en las carreras universitarias afines a la ingeniería de sistemas, donde se presentan problemas graves para desarrollar las competencias, del orden del saber hacer algoritmos.

Consecuentemente, se pretende diseñar una metodología para el diseño de algoritmos capaz de coadyuvar al desarrollo de competencias propositivas para la creación de algoritmos en la educación superior. Con ello se contribuye a identificar cuáles son los principios, las prácticas, roles, los métodos, artefactos y demás aspectos que se deben tener en cuenta para diseñar este tipo de metodología (Hurtado, 2009).

2. Fundamentos

La idea intuitiva más general de un algoritmo es un procedimiento que consiste en un conjunto finito de instrucciones que, dada una entrada de un conjunto de posibles entradas, permite obtener una salida si existe una salida tales o bien obtener nada en absoluto si hay salida para esa entrada en particular a través de una ejecución sistemática de las instrucciones.

El conjunto de posibles entradas se compone de todas las entradas a la que el algoritmo da una salida. Si hay una salida para una entrada particular, entonces se dice que el algoritmo se puede aplicar a esta entrada y procesarla para dar la salida correspondiente (Díaz, Montero & Aedo, 2005).

Se requiere que un algoritmo se detiene en cada entrada, lo que implica que cada instrucción requiere una cantidad finita de tiempo, y cada entrada tiene una longitud finita. También requerimos que la salida de una entrada legal a ser único, es decir, el algoritmo es determinista en el sentido de que el mismo conjunto de instrucciones que se ejecutan cuando se inicia el algoritmo en una entrada particular, más de una vez (Joyanes, 1998; Iglesias, 1998).

El diseño y análisis de algoritmos son de fundamental importancia en el campo de la informática (Estrada, 1996; Deitel & Deitel, 1999). Esto no debería ser sorprendente, ya que todas las áreas de ciencias de la computación dependen en gran medida el diseño de algo eficiente (Hurtado & Neira, 1996; Jaworski, 1999).

La referencia a las “instrucciones” en los algoritmos implica que hay algo o alguien capaz de entender y seguir las instrucciones dadas (Sánchez Allende, 2001; Lea, 2001). A esto le llamamos un “computador”, que se han vuelto indispensables en casi todo lo que hacemos (Díaz, Montero & Aedo, 2005). Tenga en cuenta, sin embargo, que, aunque la mayoría de los

algoritmos están efectivamente destinados a implementación informática eventual, la noción de algoritmo no depende de tal suposición (Pressman y Troya, 1988; Sommerville, 2005).

Podemos considerar algoritmos para ser soluciones a los problemas de procedimiento (Ackoff, 1998). Estas soluciones no son respuestas, pero las instrucciones específicas son determinantes para obtener respuestas (Senn, 1992). Es este énfasis en los procedimientos constructivos definidos con precisión que hace que la informática distinta de otras disciplinas. En particular, esto la distingue de las matemáticas teóricas, cuyos practicantes suelen ser satisfecho con sólo probar la existencia de una solución a un problema y, posiblemente, la investigación de las propiedades de la solución (Hurtado de Barrera, 2010).

Desde una perspectiva práctica, lo primero que hay que hacer antes de diseñar un algoritmo es entender completamente el problema dado. Se recomienda leer la descripción del problema cuidadosamente y hacer preguntas si tiene alguna duda, también, hacer unos pequeños ejemplos a mano, pensar en casos especiales, e indagar de nuevo si es necesario (Lovos & Gibelli, 2012).

Hay algunos tipos de problemas que surgen en el desarrollo de aplicaciones informáticas con bastante frecuencia, es posible que pueda utilizar un algoritmo conocido para resolverlo (Pato, 2009; Porto, 2016). Por supuesto, esto ayuda a entender cómo funciona un algoritmo tal y conocer sus fortalezas y debilidades, sobre todo si tiene que elegir entre varios algoritmos disponibles. Pero, a menudo no se encuentra un algoritmo que se ajuste completamente y se tendrá que diseñar uno propio (Cataldi, Lage, Pessacq & García-Martínez, 2003).

Por lo general, queremos que nuestros algoritmos de poseer varias cualidades. Después del funcionamiento, el más importante es la eficiencia. De hecho, hay dos clases de eficiencia algoritmo: la eficiencia del tiempo, lo que indica qué tan rápido se ejecuta el algoritmo y la eficiencia del espacio, lo que indica la cantidad de memoria extra que utiliza.

Otra característica deseable de un algoritmo es la simplicidad. A diferencia de la eficiencia, que puede definirse con precisión e investigado con rigor matemático, la sencillez, como la belleza, está en gran medida en el ojo del espectador (Sepúlveda-Aguirre et al., 2018).

Debido a que los algoritmos más simples son más fáciles de entender y fácil de programar; en consecuencia, los programas resultantes generalmente contienen menos errores. También está el atractivo estético innegable de la simplicidad (Bertoa, Troya & Vallecillo, 2002). A veces algoritmos más simples también son más eficientes que las alternativas más complicados. Desafortunadamente, no siempre es cierto, en cuyo caso un compromiso juicioso necesita ser hecho (Solano et al., 2017).

Sin embargo, otra característica deseable de un algoritmo es la generalidad. Hay, dos vías: la generalidad del problema del algoritmo resuelve y el conjunto de entradas que acepta. En el primer tema, tenga en cuenta que a veces es más fácil diseñar un algoritmo para un problema planteado en términos más generales. Hay situaciones, sin embargo, donde el diseño de un algoritmo más general es innecesario o difícil o incluso imposible.

3. Funcionamiento y fases

El ciclo de vida de la metodología propuesta se compone de cuatro fases: Requisitos, Análisis, Diseño y Validación, los cuales se muestran gráficamente en la Figura 2 y a continuación se describen cada una de ellas.

Grafica 1. *Fases de la metodología propuesta.*
Fuente: *Elaboración Propia (2016)*

3.1. Fase de requisitos

El propósito de esta fase se centra, primero, en definir el problema que se desea solucionar, en términos de identificar qué es lo que se debe automatizar para lograr un efecto de solución. En segundo término, se determinan los requisitos y pautas que le exige la problemática al algoritmo. En cuanto a estos requisitos que representan las necesidades a resolver, las restricciones, las precisiones y salvedades, y todas aquellas características que debe poseer el algoritmo a ser diseñado.

Los entregables en esta etapa corresponden a una especificación de los requisitos (Formato en la Figura 2) y de la descripción de la problemática, destacando las causas, síntomas y signos y las consecuencias (Formato en la Figura 1).

Problema a solucionar	
Descripción:	
Causas:	
Síntomas y signos:	
Consecuencias:	

Figura 1. *Formato de entregable del problema a solucionar*
Fuente: *Elaboración propia.*

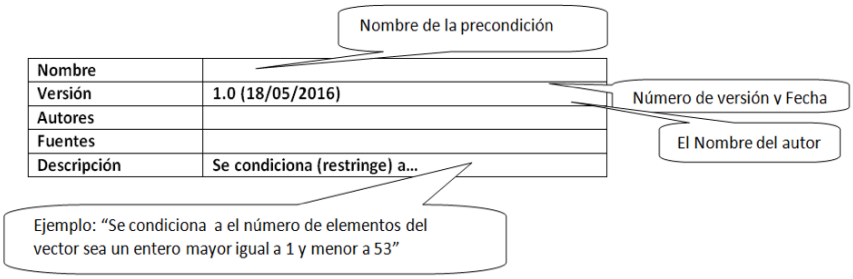
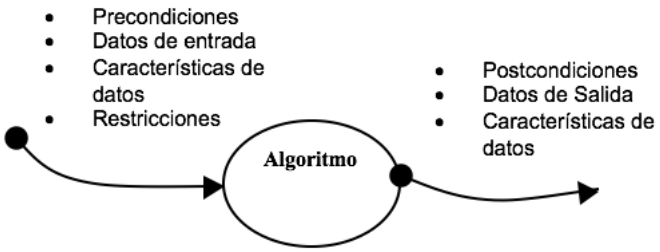


Figura 2. *Formato de entregable de los requisitos*
Fuente: *Elaboración propia.*

3.2. Fase de análisis

En la etapa de análisis se examinan y se especifican las precondiciones –condiciones de inicio y condiciones de los datos de entrada para el funcionamiento y cumplimiento de los requisitos del algoritmo–; las postcondiciones o condiciones que deben cumplirse como producto de la ejecución del algoritmo; los datos de entrada y de salida y sus características;

y finalmente, las restricciones propias de la problemática. En la Grafica 3, se muestra claramente un esquema de lo que se busca en la fase de análisis.



Grafica 2. *Esquema de la fase de análisis.*
Fuente: *Elaboración Propia (2016)*

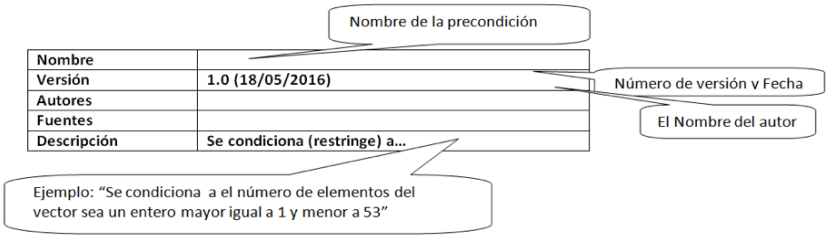


Figura 3. *Formato de entregable de las precondiciones y restricciones*
Fuente: *Elaboración propia.*

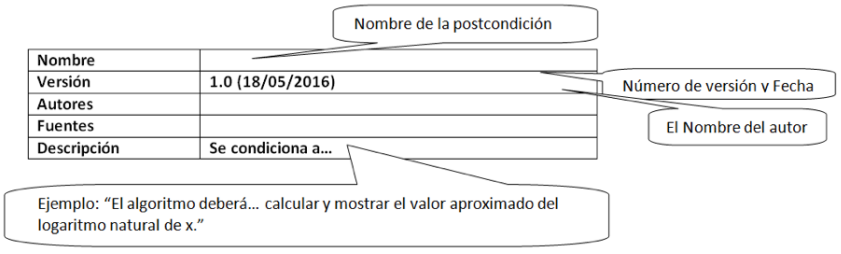


Figura 4. *Formato de entregable de las postcondiciones*
Fuente: *Elaboración propia.*

3.3. Fase de diseño

En esta fase se utilizan las distintas primitivas algorítmicas en tres periodos o estadios

- Captura de datos e inicialización. Se realizan los procedimientos para asegurar el cumplimiento de las precondiciones.
- Proceso de resolución de requisitos. Se utiliza la creatividad y las distintas primitivas para dar cumplimiento a cada uno de los

requisitos que le son encargados al algoritmo

- Presentación de datos. Se verifican el logro y la realización de las postcondiciones.

Grafica 3. *Periodos de la fase de diseño.*
Fuente: *Elaboración propia (2016)*

3.4. Fase de validación

En esta etapa se busca validar el cumplimiento de todos los posibles casos de los requisitos, así como el de ajustar el funcionamiento del algoritmo para dar cumplimiento a las postcondiciones desde la perspectiva de las precondiciones. La validación utiliza la técnica de la prueba de escritorio, que no es más que efectuar un proceso de simulación del funcionamiento algoritmo desarrollado, probando cada uno de los casos colocando los valores tipos que deben representar al conjunto de entradas y ejecutando cada instrucción para verificar el cumplimiento de los requisitos y de las postcondiciones. En el caso de encontrar desviaciones de debe corregir el algoritmo y realizar la prueba de escritorio nuevamente.

4. Conclusiones

El primer objetivo específico, se cumple al analizar los modelos, prácticas, recursos y artefactos que se deben considerar al realizar un ciclo de vida de una metodología que se tomaron sobre la base para desarrollar la metodología propuesta.

En lo que atañe al objetivo específico de los factores de éxito para poder construir algoritmos que respondan a las necesidades de formación de las instituciones de educación superior, se alcanzó por medio del establecimiento la especificación de la metodología para la construcción de algoritmos.

Con el diseño y presentación de las fases de la metodología, se da cumplimiento al objetivo que plantea proponer un ciclo de vida de la metodología capaz de generar algoritmos que sea aplicable en los procesos de enseñanza aprendizaje en la educación superior. A partir del desarrollo de la metodología propuesta, se derivan nuevos siguientes interrogantes:

¿Cómo evaluar una metodología para la construcción de algoritmos desarrollada, con la que se pretende mejorar el proceso de enseñanza aprendizaje del diseño de algoritmos?

¿Cómo sería el proceso de implementación de la metodología en el proceso de enseñanza aprendizaje?

Referencias Bibliográficas

Ackoff, R. (1998). El Arte de Resolver problemas, Noriega Editores- Decimocuarta reimpresión.

Aguilar, J., Muñoz, J. & Pomares, S. (2010). Guías de diseño para el desarrollo de objetos de aprendizaje. Obtenido de http://ingsw.ccbas.uaa.mx/sitio/images/publicaciones/ENC04(Aguilar_Munoz).pdf

Alba, E. (2009). Análisis y diseño de algoritmos genéticos paralelos distribuidos. Tesis de grado- Universidad de Málaga. Málaga, España: Universidad de Málaga.

Arsham, H. (1995). Interactive education: Impact of the internet on learning & teaching. [On-Line], Aviliable:http://UBMAIL.ubalt.edu/i«harsham/interactive.htm

Bertoa, M., Troya, J. & Vallecillo, A. (2002). Aspectos de Calidad en el Desarrollo de Software Basado en Componentes. Depto. Lenguajes y Ciencias de la Computación. Universidad de Málaga. [On-Line], Aviliable: http://www.lcc.uma.es/~av/Publicaciones/02/CalidadDSBC.pdf

Cataldi, Z., Lage, F., Pessacq, R. & García-Martínez, R. (2003). Metodología extendida para la creación de software educativo desde una visión integradora. RELATEC.

Cohoon, J. P., & Davidson, J. W. (2002). C++ program design: an introduction to programming and object-oriented design.

Deitel, H. M. & Deitel, P. J. (1999). Cómo programar en Java. 3º Ed. Prentice-Hall.

Díaz, M, Montero, S. & Aedo, I. (2005) Ingeniería Web y patrones de diseño. Universidad Carlos III Madrid. Prentice – Hall, Madrid. 409 p.

Estrada, E. L. (1996). Teoría general de sistemas aplicada a la solución integral de problemas. Universidad del Valle.

Hurtado de Barrera, J. (2010). Metodología de la investigación. Guía para la comprensión holística de la ciencia (Cuarta ed.). Caracas: Quirón Ediciones.

Hurtado, D. & Neira, M. (1996). Software aplicativo a la enseñanza de la asignatura sistemas operacionales. Tesis de Grado. Universidad del Norte. Barranquilla.

Hurtado, D. (2009). Metodología para el desarrollo de sistemas basados en objetos de aprendizaje: MethSysOL (Master's thesis, Maestría en Ingeniería de Sistemas).

Iglesias, C. (1998). Definición de una metodología para el desarrollo de sistemas multiagente. Universidad Politécnica de Madrid, España

Jaworski, J. (1999). Java 1.2 al descubierto. Prentice–hall Madrid. 1344 p.

Joyanes, L. & Zahonero, I. (2002). Programación en Java 2: Algoritmos, estructuras de datos y programación orientada a objetos. McGraw-Hill. Madrid. p.725

Joyanes, L. (1998). Programación Orientada a Objetos. 2° Ed. Osborne McGraw-Hill. p.895

Lea, D. (2001). Programación concurrente en Java: Principios y patrones de diseño. 2° Ed. Addison Wesley. p.430

Lovos, E. & Gibelli, T. I. (2012). Análisis de eficiencia de algoritmos. TE & ET.

Main, M. & Savitch, P. (2001). Data structures and other objects using C++. Addison Wesley. 783 p

Pato, A. S. (2009). Hacia el algoritmo. Cultura, Ciencia y Deporte, 4(10), 5.

Porto, R. S. (2016). Software móvil para la enseñanza de las Normas Internacionales de Información Financiera (NIC-NIIF) APPIFRS (parte 2). Journal of Engineering and Technology, 5(1).

Pressman, R. S. & Troya, J. M. (1988). Ingeniería del software.

Sánchez Allende, J. (2001). Java 2 iniciación y referencia.

Selic, B. Gullekson, G. and Ward, P. (1994) Real-Time Object-Oriented Modeling. John Wiley & Sons. 525 p.

Senn, J. (1992). Análisis y diseño de sistemas de información. 2º ed. McGraw-Hill. México. 914 p

Sepúlveda-Aguirre, J., López, D. M., Solano, R. P., Marchena, L. M., Gutiérrez, C. A., & Giraldo, L. F. (2018). El sector público como actor relevante de la ciencia, la tecnología y la innovación en Colombia.

Solano, R. P., Sarmiento, A. S., Giraldo, L. F. G., Solano, A. P., & López, D. S. (2017). MARCO DE ENSEÑANZA BASADO EN OVAs PARA EL ADIESTRAMIENTO EN LA INTRODUCCIÓN DE LOS ALGORITMOS (OLIA). Revista científica electrónica de Educación y Comunicación en la Sociedad del Conocimiento, 17(1).

Sommerville, I. (2005). Ingeniería del Software: Un Enfoque Práctico. Eddison Wesley, México.

DISEÑO DE UN MODELO PARA ASISTENCIA MÉDICA Y PREDICCIÓN DE PREECLAMPSIA EN MUJERES EMBARAZADAS•

Roberto Porto Solano *
Miguel Jimeno Paba **
Eduardo Zurék Varela ***
Francisco Hernández Caballero****

• Capítulo de libro resultado de investigación. Producto Colaborativo de los grupos de investigación Aglaia y GReCIS. Producto resultado de tesis de maestría "Diseño de un modelo para asistencia médica y predicción de preeclampsia en mujeres embarazadas" (2017).

* Ingeniero de Sistemas. Magister en ingeniería de Sistemas y Computación. Corporación Universitaria Americana Correo: rporto@coruniamericana.edu.co
** Ingeniero de Sistemas. Magister en ingeniería de Sistemas y Computación. Doctor en ciencias de la computación y la ingeniería. Universidad del Norte. Correo: majimeno@uninorte.edu.co
*** Ingeniero de Sistemas. Magister en ingeniería eléctrica. Doctor en filosofía. Universidad del Norte Correo: ezurek@uninorte.edu.co
**** Contador público. Magister en economía financiera. Corporación Universitaria Americana Correo: fhernandez@coruniamericana.edu.co

Resumen

Objetivo presentar un modelo basado de un sistema de información para la asistencia de mujeres embarazadas, que les ayudara a tener un mejor control de su embarazo y para el médico poder estar mejor informado sobre las probabilidades de que sus pacientes puedan padecer de preeclampsia.

Métodos Dentro de esta investigación se desarrolló un modelo de predicción de preeclampsia en mujeres embarazadas durante el segundo trimestre de gestación, empleando técnicas de minería de datos (Arboles de decisión, Algoritmo C5.0, Principal Component Analysis, Non Metric Dimensional Scaling). El modelo surge del interés de diseñar una herramienta de predicción de preeclampsia, con el fin de ayudar con los expertos en obstetricia a examinar más fácilmente los resultados obtenidos de sus pacientes apoyando la toma de decisiones. Se determinaron las variables para hacer la predicción, para proceder a seleccionar el mejor método para lograr este proceso.
Resultado el modelo fue probado sobre un conjunto de 100 datos simulados de los cuales el 75% de la muestra, corresponde a datos de entrenamiento y el 25 %, a datos de validación. Los resultados obtenidos de las técnicas (C5.0, PCA, NMDS) se compararon con otras técnicas de predicción encontradas en revisión literaria mejorando el porcentaje de acierto.
Conclusión el sistema de clasificación de preeclampsia en mujeres embarazadas desarrollado, muestra cómo pueden implementar técnicas de inteligencia artificial en medicina enfocadas a las áreas ginecología, para el reconocimiento automático de enfermedades en las entidades de atención médicas de tercer nivel.
Palabras Claves: Predicción, Preeclampsia, minería de datos, Análisis multivariado, c5.0.

Abstract

Objective To present a model based on an information system for the care of pregnant women, which will help them to have a better control of their pregnancy and for the doctor to be better informed about the chances that their patients may have preeclampsia.
Methods A predictive model of preeclampsia was developed in pregnant women during the second trimester of pregnancy, using data mining techniques (Decision Trees, C5.0 Algorithm, Principal Component Analysis, Non Metric Dimensional Scaling). The model emerges from the interest of designing a preeclampsia prediction tool to help obstetrics experts to more easily examine the outcomes of their patients by supporting decision making. The variables were determined to make the prediction, to proceed to select the best method to achieve this process.

Result The model was tested on a set of 100 simulated data of which 75% of the sample corresponds to training data and 25% to validation data. The results obtained from the techniques (C5.0, PCA, NMDS) were compared with other prediction techniques found in literary revision, improving the percentage of success.
Conclusion The preeclampsia classification system in pregnant women developed shows how they can implement artificial intelligence techniques in medicine focused on gynecological areas, for automatic disease recognition in third-level medical care entities.

Keywords: Prediction, Preeclampsia, data mining, multivariate analysis, C5.0.

Introducción

La preeclampsia es el trastorno hipertensivo más frecuente en el embarazo y constituye la primera causa de mortalidad materna en nuestro medio (Cifuentes, 2006). La incidencia mundial es de 2-10% (Harskamp, 2007). En E.U.A. afecta al 3-8% de las mujeres en embarazo, es la segunda causa de muerte materna luego del embolismo representando aproximadamente un 15 % de esas muertes (Sibai, 2005). Así mismo es la primera causa de prematuridad iatrogénica en los países industrializados. En Colombia afecta aproximadamente al 10-15% de la población (Cifuentes, 2010).

Más recientemente se ha encontrado que la disfunción endotelial de la preeclampsia se asocia también a un incremento en la producción de factores antiangiogénicos: sFlt-1 (forma soluble de tirosina kinasa) y sEng (endoglina Soluble). Estos antagonizan los factores antigénicos: VEGF (factor de crecimiento del endotelio vascular) y PIGF (factor de crecimiento placentario), creándose un balance negativo de estos últimos. En general se considera que una paciente es hipertensa en el embarazo cuando tiene una presión arterial igual o mayor que 140 y /o 90mmHg (Cifuentes, 2009).

El problema de la detección de preeclampsia radica en el análisis de síntomas que permitan conocer el comportamiento (estado) de un proceso gestacional, con el fin de detectar anomalías. Según Vázquez (2011), a menudo las mujeres que padecen preeclampsia no se sienten enfermas, ya que algunos síntomas iniciales en el embarazo suelen confundirse con trastornos normales del embarazo.

Según Bautista (2012), los trastornos hipertensivos asociados al embarazo son muy frecuentes durante el control prenatal. En nuestro medio es una entidad endémica presente hasta en 10% de los embarazos. En la práctica diaria esta prevalencia tan importante se acompaña de gran morbilidad y mortalidad materna y perinatal. La primera causa de muerte materna en nuestro país está relacionada con la toxemia gravídica. La entidad obstétrica más relacionada con estos casos letales es la eclampsia, acompañada o agravada por el Síndrome de Hemolytic anemia Elevated Liver Enzyme Low Platelet HELLP. La manifestación hipertensiva más frecuentemente encontrada durante el embarazo es la preeclampsia. Esta entidad aparece en gestaciones menores de 34 semanas en 35% de los casos. Cuando se encuentra la asociación de preeclampsia y prematurez

se conjuga el determinante clínico de 30% de las muertes perinatales en nuestro medio. Es así como la preeclampsia es la segunda causa de muerte perinatal. A nivel mundial se ha intentado controlar de manera preventiva la aparición de los cuadros clínicos severos relacionados a la hipertensión arterial durante la gestación sin lograr el éxito esperado. Estas entidades están muy relacionadas con las condiciones socioeconómicas de la población, lo cual hace imposible su desaparición del escenario clínico. Su frecuencia se dispara en casos de madres solteras, embarazos no deseados, desempleados y desplazados, toda ésta una problemática muy nuestra. Sin embargo, el pronóstico es susceptible de modularse mejorando el diagnóstico temprano.

Según la investigación realizada en el centro de investigación Harris Birthright de Medicina Fetal en Londres por Neocleous (2009). De 6838 casos de mujeres embarazadas se registraron por paciente 24 parámetros. De los cuales 15 parámetros fueron considerados, como los de más influencia en la caracterización del riesgo de aparición de preeclampsia. Para la predicción de la preeclampsia fue utilizado una red neuronal feedforward tanto multicapa como multi-losa. En el entrenamiento de la red neuronal lograron una clasificación de los casos del 83,6% de la preeclampsia y en una prueba ajustada de un 93,8%. Los casos de predicción de preeclampsia cuando apareció información de tabaquismo y alcoholismo llegaron las pruebas a un 100%.

Según Dustin y Dunsmuir (2014), el modelo fullPIERS (Preeclampsia Intégrate Estímate of RiSk) fue desarrollado para evitar la necesidad de no usar pruebas de laboratorio que con frecuencia son realizadas en hospitales de tercer nivel. Para la efectividad de predicción se toman variables como: simple demografía, los síntomas y signos clínicos; Este modelo de predicción simplificado incluye la paridad, la edad gestacional en la presentación, la presión arterial sistólica, diastólica, la varilla de medición proteinuria, y la presencia de los siguientes síntomas: dolor en el pecho, falta de aliento, dolor de cabeza, trastornos visuales, y sangrado vaginal con dolor abdominal. El desarrollo de la aplicación móvil para el diagnóstico y la gestión de las mujeres embarazadas con preeclampsia es descrito. Esta aplicación está diseñada para su uso por basado en la comunidad proveedores de atención médica (c-SCPH) en los centros de salud y durante visitas a domicilio para recoger síntomas y realizar mediciones clínicas (incluyendo lecturas del oxímetro de pulso). Los datos clínicos recogidos en las mujeres con preeclampsia se utilizan como

entradas a un modelo predictivo, que proporciona una puntuación de riesgo para el desarrollo de resultados adversos. Sobre la base de este riesgo, la aplicación proporciona recomendaciones sobre el tratamiento, remisión, y la reevaluación. C-HCP puede acceder registros de los pacientes a través de múltiples visitas, utilizando múltiples dispositivos que se sincronizan mediante una captura de Investigación electrónico de datos segura servidor. Una característica única de esta aplicación es la capacidad de medir la saturación de oxígeno con un pulsioxímetro conectado a un teléfono inteligente (oxímetro de teléfono). El modelo miniPIERS emite una puntuación, una calificación de probabilidad, la cual es utilizada dentro de un modelo de árbol de decisión para dar las recomendaciones a los tratamientos de los pacientes. Además, el árbol de decisión incluye información de los eventos maternos adversos en el pasado, con la actual etapa del embarazo, y SpO2.

En S.Y. Leemaqz (2013), obtienen los valores influyentes de los exámenes (NPV, PPV, r, s) para la predicción de preeclampsia antes de la semana 15 y dentro de la semana 15 a la 20. Donde NPV es la proporción de verdad negativos en los casos previstos de embarazo sin complicaciones, PPV es la proporción de verdad positivos en los casos previstos de PE, r es la proporción de casos verdaderamente predichos de PE y s es la proporción de verdaderamente casos de embarazo sin complicaciones predicho. El modelo expuesto utiliza un algoritmo N.Bayes para predecir la preeclampsia basado en los valores NPV, PPV, r, s obteniendo un porcentaje de aciertos de 81%.

Según Etchegaray Adolfo (2012), las características maternas más influyentes para la predicción de preeclampsia se pueden obtener entre las semanas 11 y 13 o sea en el segundo trimestre de gestación. La técnica para la predicción es basada en la definida por el grupo de la Fetal Medicine Foundation de reino unido, en donde hacen una combinación de factores de riesgos maternos como (IMC, grupo étnico antecedentes de preeclampsia, de presión arterial, combinación de analitos séricos, índice de pulsatilidad de arterias uterinas, proteínas asociadas al embarazo y factor de crecimiento placentario).

Según Garovic (2014), la utilización de mediciones de orina y sueros circulantes de marcadores angiogénicos no proporcionan una forma fiable para la detección de preeclampsia. En un intento por mejorar los valores predictivos de los marcadores se combinaron índices angiogénicos como

sFlt-1/PIGF y encontraron cambios que demuestran que no es efectiva en la predicción de la preeclampsia de inicio tardío, pero que es útil en la predicción temprana de la enfermedad. Estos resultados plantean la posibilidad de utilizar estas pruebas en los subgrupos de pacientes especiales, tales como aquellos con hipertensión crónica, que puede estar en un nivel superior de riesgo de preeclampsia y las complicaciones relacionadas con preeclampsia.

También Verlohren (2012), explica que en numerosos estudios que se han puesto de manifiesto, se denota el papel clave de la placenta en la patogénesis de la preeclampsia. Un cambio en la sFlt-1 (similar a Fms solubles tirosina quinasa-1) / PlGF (factor de crecimiento placentario) relación se asocia con la enfermedad. Aunque la preeclampsia parece ser una enfermedad claramente definida, la presentación clínica, y en particular la dinámica de la evolución clínica, pueden variar enormemente. Las únicas herramientas disponibles a diagnosticar la preeclampsia son la medición de la presión arterial y el muestreo de proteína en la orina. Sin embargo, estas herramientas tienen una baja sensibilidad y especificidad con respecto a la predicción del curso de la enfermedad o los resultados maternos y perinatales. La única cura para la enfermedad es la entrega, aunque un diagnóstico oportuno ayuda en la disminución de la morbilidad y mortalidad materna y fetal. Los índices angiogénicos sFlt1 / PlGF en proporción son capaces de dar una valiosa información adicional sobre el estado y evolución de la enfermedad y es apto para ser implementado en el algoritmo diagnóstico de la preeclampsia.

Por otro lado, Loaiza, Solano, Simancas y Ojito (2017), proponen varios modelos para estimar la demanda de carga aérea en Colombia, obtenidos mediante el uso de metodologías como la regresión lineal y las redes neuronales, que se pueden usar para caracterizar la demanda actual y para pronosticar escenarios de demanda futura en ciertos contextos. Con el fin de ajustar la estimación de los modelos presentados, se utilizó la demanda de información aérea de los 19 aeropuertos más importantes del país (en términos de carga enviada), registrada por la Unidad Administrativa Especial de Aeronáutica Civil en Colombia (Aerocivil) de 2005 a 2014, además de información socioeconómica sobre las áreas de influencia de dichos aeropuertos, obtenida principalmente por el Departamento Administrativo de Estadísticas Nacionales (DANE) del mismo país. Finalmente, se establece una comparación entre los resultados obtenidos por cada metodología de modelado, encontrando mejores resultados con los modelos de redes neuronales.

Metodología

En este artículo se presentan los aspectos metodológicos y técnicas usadas para la obtención de la información. Además, se delimita el objeto de estudio de este proyecto y los procedimientos utilizados para poder llevar a cabo la investigación.

La investigación está compuesta por un método de investigación cuantitativo con componentes cualitativas, explicado a continuación:

- Análisis cuantitativo, las pruebas estadísticas de predicción en el modelo de la preeclampsia durante el segundo trimestre de gestación propuesto, se posibilitaron mediante el estudio de dos conjuntos de datos a) mujeres embarazadas que presentaron preeclampsia y b) las que no presentaron, con variables cualitativas se les dio un peso cuantitativo para, hacer más eficiente el proceso de correlación cuantitativa (otras variables), fundamental para determinar el nivel de eficiencia del método de predicción.

Modelo de predicción de preeclampsia

El modelo de detección de preeclampsia, basado en clasificación variables predictoras usando los algoritmos C5.0 evolución del C4.5 de Quinlan (2014) y NMDS que ayudaran a crear un conjunto de reglas de clasificación, provenientes del entrenamiento con el fin de caracterizar patrones de detección de preeclampsia que ayuden al experto en obstetricia a examinar y ver0ificar más fácilmente los resultados obtenidos para apoyar la toma de decisiones. Luego se hace la descripción del modelo propuesto de detección, el cual no pretende cubrir todos los aspectos involucrados en un proceso KDD, sino solo aquellos que permitan representar el conocimiento a extraer en términos de reglas de clasificación a partir de un árbol de decisión como lo explican Cook y Goldman (1984). De esta manera, el modelo presenta un conjunto de fases que permiten la preparación de datos, extracción, predicción o clasificación y presentación del conocimiento, comunes en cualquier proceso de minería de datos.

3.1. Modelos de extracción de conocimiento para la predicción de preeclampsia

En este trabajo se plantea un modelo de detección de preeclampsia, basado en análisis de datos utilizando técnicas de minería de datos, específicamente para un conjunto de datos de varias entidades prestadoras de servicios de salud (hospitales). El modelo se presenta en la Figura1, la cual presenta a grandes rasgos la iteración de diferentes módulos y la comunicación con elementos externos al sistema (Archivos). El modelo se compone de varios módulos que proporcionan una funcionalidad específica, con el fin de obtener conocimiento de la relación que existe entre el conjunto de datos. Los módulos mencionados son:

1. Preparación y selección de datos: Encargado de la recolección de la fuente de información, y preparación del conjunto de datos para la extracción del conocimiento.
2. Extracción del conocimiento. Es el impulsor del modelo de detección propuesto, y está conformado por:

    - Algoritmo de Aprendizaje. Para el proceso se construye un árbol de decisión a través del algoritmo C5.0 generando un conjunto de reglas de clasificación. El C5.0 divide la muestra en función del campo que ofrece la mayor ganancia de información. Las distintas submuestras definidas por la primera división se vuelven a dividir, por lo general basándose en otro campo, y el proceso se repite hasta que resulta imposible dividir las submuestras de nuevo. Por último, se vuelven a examinar las divisiones del nivel inferior, y se eliminan o podan las que no contribuyen significativamente con el valor del modelo.
    - Eliminación de factores. Para obtener la mayor ganancia y efectividad dentro del modelo, se eliminan los factores que no aportaron ganancia de información dentro del modelo.

3. Predicción: Encargado de inferir una predicción sobre un conjunto de datos nuevos. Para ello se parte del conocimiento encontrado con el algoritmo C5.0 para hacer la clasificación del conjunto de datos. Luego con la técnica PCA y PCOA se observa la separación de los datos y posteriormente se comparan los resultados con los obtenidos por el algoritmo C5.0 para la toma de decisión.

4. Presentación de la información final al doctor: se muestran los valores obtenidos de la predicción como los presentados en la presentación de prescripciones médicas (Porto Solano et al., 2017)

y en los sistemas de pronósticos agrícolas (García, Porto, Simancas, Molina & Hernández, 2016). Los pasos se muestran en la Figura 1. Modelo de extracción de conocimiento para la detección de Preeclampsia.

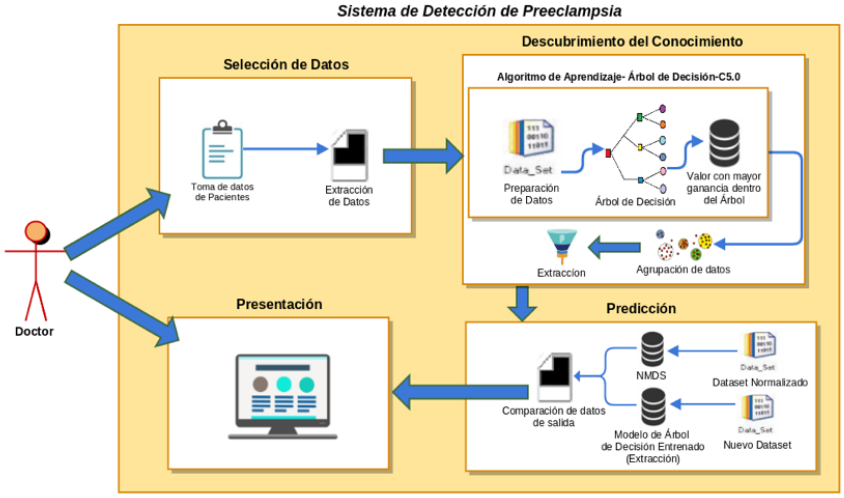


Figura 1. *Modelo de extracción de conocimiento para la detección de Preeclampsia*
Fuente: *Elaboración del autor.*

3.2. Preparación y selección de datos

En esta fase se determinan, seleccionan y extraen los datos de la fuente de información para el proceso de extracción de conocimiento. En esta etapa se realizan las siguientes actividades.

1. Extracción de datos relevantes de la fuente de información haciendo énfasis en:

   - Identificación de los datos relevantes en las fuentes de información (para el caso, un modelo relacional de tablas).
   - Extracción e integración. A pesar de que los aspectos importantes a evaluar del modelo pueden estar distribuidos en diferentes tablas en una misma base de datos, en el modelo propuesto, se considera que estos datos conforman una sola tabla, a partir de una consulta en la base de datos que retorna una tabla virtual (conjunto de registros), con los datos a utilizar en el proceso de extracción de conocimiento.

2. Preparación de datos, en esta parte se hace énfasis en:
   - Eliminación de registros inconsistentes.
   - Eliminación de atributos que no aportan información, por ejemplo, redundancia de información, valores perdidos, etc.
   - Ruido e inconsistencia de datos.
   - Desratización de atributos.

3. Selección de la información a utilizar en el proceso de extracción de conocimiento, a partir de determinar cuáles van a hacer las variables independientes y cuya variable será objetivo (conclusión). Para esto, el usuario es el encargado de determinar las variables a considerar en el proceso de aprendizaje y extracción de reglas.
4. Presentación de los resultados del aprendizaje. Esta fase puede ser soportada por herramientas estadísticas, herramientas de limpieza y otras, de tal manera que el usuario pueda realizar estas operaciones de una forma eficiente.

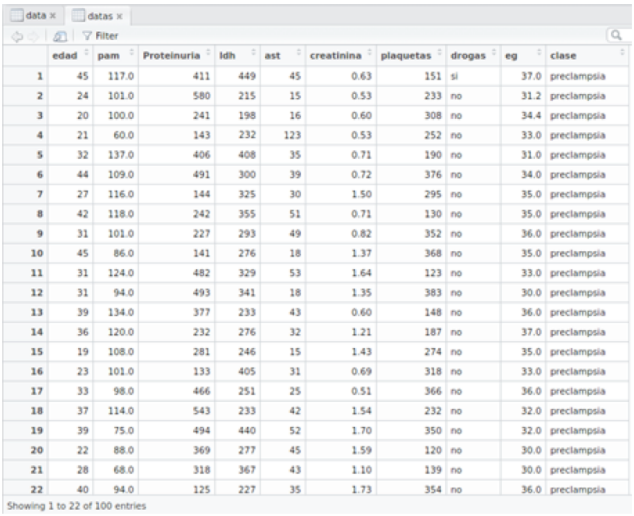
4. Implementación del modelo

En esta etapa se presenta la implementación del modelo de detección de preeclampsia, basado en clasificación variables predictoras usando el algoritmo C5.0 y NMDS y en el cual podemos observar la estructura de la información ingresada, aplicabilidad de los algoritmos, resultados y graficas de resultados.

4.1. Algoritmo C5.0 en R

1. Se leen los datos y se almacena en la variable datas y se imprime

```
require(gdata)
library(gdata)
read.xls(/home/topor/Descargas/prueba.xls)
datas= read.xls(/home/topor/Descargas/prueba.xls)
```



data × datas ×

Filter

	edad	pam	Proteinuria	ldh	ast	creatinina	plaquetas	drogas	eg	clase
1	45	117.0	411	449	45	0.63	151	si	37.0	preclampsia
2	24	101.0	580	215	15	0.53	233	no	31.2	preclampsia
3	20	100.0	241	198	16	0.60	308	no	34.4	preclampsia
4	21	60.0	143	232	123	0.53	252	no	33.0	preclampsia
5	32	137.0	406	408	35	0.71	190	no	31.0	preclampsia
6	44	109.0	491	300	39	0.72	376	no	34.0	preclampsia
7	27	116.0	144	325	30	1.50	295	no	35.0	preclampsia
8	42	118.0	242	355	51	0.71	130	no	35.0	preclampsia
9	31	101.0	227	293	49	0.82	352	no	36.0	preclampsia
10	45	86.0	141	276	18	1.37	368	no	35.0	preclampsia
11	31	124.0	482	329	53	1.64	123	no	33.0	preclampsia
12	31	94.0	493	341	18	1.35	383	no	30.0	preclampsia
13	39	134.0	377	233	43	0.60	148	no	36.0	preclampsia
14	36	120.0	232	276	32	1.21	187	no	37.0	preclampsia
15	19	108.0	281	246	15	1.43	274	no	35.0	preclampsia
16	23	101.0	133	405	31	0.69	318	no	33.0	preclampsia
17	33	98.0	466	251	25	0.51	366	no	36.0	preclampsia
18	37	114.0	543	233	42	1.54	232	no	32.0	preclampsia
19	39	75.0	494	440	52	1.70	350	no	32.0	preclampsia
20	22	88.0	369	277	45	1.59	120	no	30.0	preclampsia
21	28	68.0	318	367	43	1.10	139	no	30.0	preclampsia
22	40	94.0	125	227	35	1.73	354	no	36.0	preclampsia

Showing 1 to 22 of 100 entries

Figura 2. *Dataset Preeclampsia*
Fuente: *Elaboración del autor.*

2. Vemos los primeros datos de la variable del dataset como se observa en la Figura 3.

```
head(datas)
```

```
  edad pam Proteinuria ldh ast creatinina plaquetas drogas   eg       clase
1   45 117         411 449  45       0.63       151     si 37.0 preclampsia
2   24 101         580 215  15       0.53       233     no 31.2 preclampsia
3   20 100         241 198  16       0.60       308     no 34.4 preclampsia
4   21  60         143   9 123       0.53       252     no 33.0 preclampsia
5   32 137         406 408  35       0.71       190     no 31.0 preclampsia
6   44 109         491 300  39       0.72       376     no 34.0 preclampsia
> |
```

Figura 3. *Encabezado del Dataset*
Fuente: *Elaboración del autor.*

3. Se muestra la estructura de la información almacenada como se observa en la Figura 4.

```
str(datas)
```

```
'data.frame':   100 obs. of  10 variables:
 $ edad       : int  45 24 20 21 32 44 27 42 31 45 ...
 $ pam        : num  117 101 100 60 137 109 116 118 101 86 ...
 $ Proteinuria: int  411 580 241 143 406 491 144 242 227 141 ...
 $ ldh        : int  449 215 198 9 408 300 325 355 293 276 ...
 $ ast        : int  45 15 16 123 35 39 30 51 49 18 ...
 $ creatinina : num  0.63 0.53 0.6 0.53 0.71 0.72 1.5 0.71 0.82 1.37 ...
 $ plaquetas  : int  151 233 308 252 190 376 295 130 352 368 ...
 $ drogas     : Factor w/ 2 levels "no","si": 2 1 1 1 1 1 1 1 1 1 ...
 $ eg         : num  37 31.2 34.4 33 31 34 35 35 36 35 ...
 $ clase      : Factor w/ 2 levels "preclampsia",..: 1 1 1 1 1 1 1 1 1 1 ...
```

Figura 4. *Estructura del Dataset*
Fuente: *Elaboración del autor.*

4. Se utilizan los factores de clasificación cruzada para construir una tabla de recuento de la variable clase en cada combinación de los niveles de los factores

```
table(datas$clase)
```

```
preclampsia sinpreeclampsia
         50              50
```

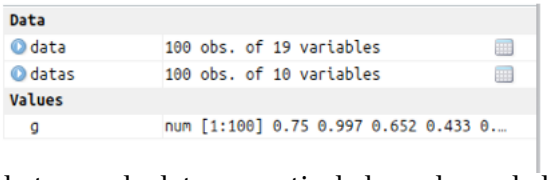
Figura 5. *Factores de clasificación cruzada de Dataset.*
Fuente: *Elaboración del autor.*

5. Generamos números aleatorios para asegurar todos los posibles resultados.

```
set.seed(9850)
```

6. Se generan valores de una distribución uniforme a partir de los factores

```
g <- runif(nrow(datas))
```



7. Se ordena la trama de datos a partir de los valores de la distribución generados

```
datar <- datas[order(g),]
```

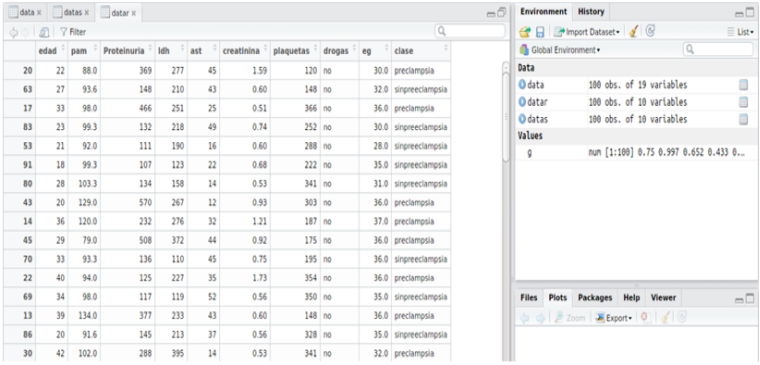


Figura 7. *Datos a partir de los valores de la distribución de Dataset.*
Fuente: *Elaboración del autor.*

Se realiza la clasificación en 80 registros del dataset a partir de los datos ordenados de la distribución para cada factor.

```
m1 <- C5.0(datar[1:80,-10], datar[1:80,10])
```

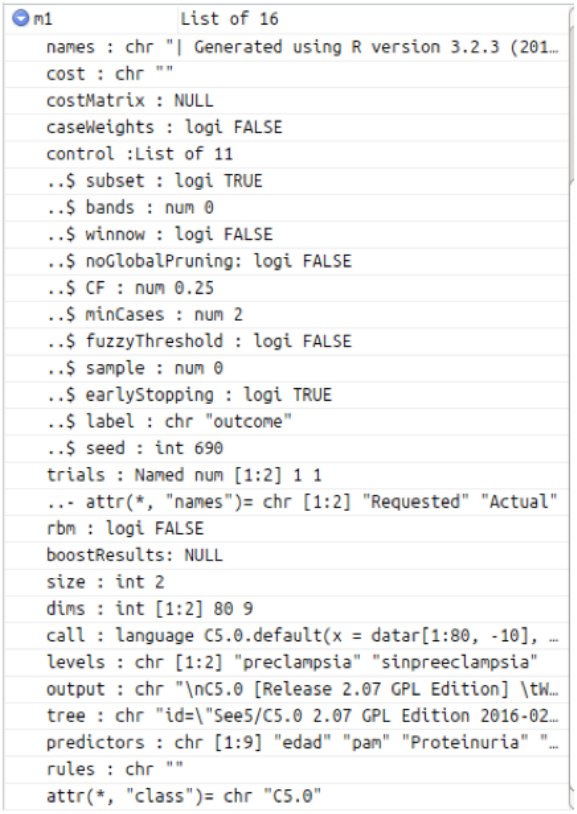


Figura 8. *Distribución para cada factor generado de Dataset.*
Fuente: *Elaboración del autor.*

9. Se imprime el resultado del entrenamiento del algoritmo c5.0

m1

```
Call:
C5.0.default(x = datar[1:80, -10], y = datar[1:80, 10])

Classification Tree
Number of samples: 80
Number of predictors: 9

Tree size: 2

Non-standard options: attempt to group attributes
```

Figura 9. *Resultado de Dataset.*
Fuente: *Elaboración del autor.*

10. Se produce el resumen de los resultados de la clasificación, mostrando el nivel del árbol generado en la clasificación

```
summary(m1)
```

```
Console ~/

Call:
C5.0.default(x = datar[1:80, -10], y = datar[1:80, 10])


C5.0 [Release 2.07 GPL Edition]        Wed Feb 10 13:08:21 2016
-------------------------------

Class specified by attribute `outcome'

Read 80 cases (10 attributes) from undefined.data

Decision tree:

ldh <= 218: sinpreeclampsia (44/2)
ldh > 218: preclampsia (36)


Evaluation on training data (80 cases):

          Decision Tree
        ----------------
        Size      Errors

           2    2( 2.5%)   <<


         (a)   (b)    <-classified as
        ----  ----
          36     2    (a): class preclampsia
                42    (b): class sinpreeclampsia


       Attribute usage:

       100.00% ldh


Time: 0.0 secs
```

Figura 10. *Resumen de los resultados de la clasificación.*
Fuente: *Elaboración del autor.*

11. Se predice la clase a través del árbol generado por el algoritmo C5.0

```
p1<-predict(m1, datar[80:100,])
   p1
```

```
 [1] sinpreeclampsia sinpreeclampsia sinpreeclampsia preclampsia     preclampsia     preclampsia     preclampsia
 [8] sinpreeclampsia preclampsia     sinpreeclampsia preclampsia     sinpreeclampsia preclampsia     sinpreeclampsia
[15] preclampsia     preclampsia     sinpreeclampsia preclampsia     preclampsia     sinpreeclampsia sinpreeclampsia
Levels: preclampsia sinpreeclampsia
```

Figura 11. *Predicción de clase.*
Fuente: *Elaboración del autor.*

12. Construye una tabla de recuento de la variable clase en cada combinación de los niveles de los factores de p1

```
table(datar[80:100,10], Predicted= p1)
```

	Predicted	
	preclampsia	sinpreeclampsia
preclampsia	11	1
sinpreeclampsia	0	9

Figura 12. *Tabla de recuento de la variable.*
Fuente: *Elaboración del autor.*

13. Se imprime el árbol generado en la clasificación como se observa en la Figura 13.

```
plot(m1)
```

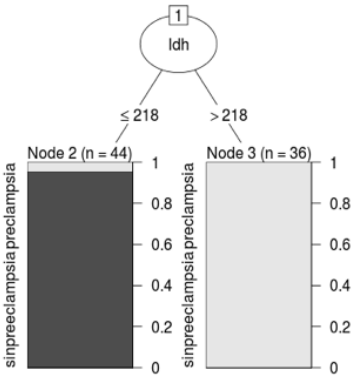


Figura 13. *Árbol entrenado C5.0*
Fuente: *Elaboración del autor.*

4.2 Sistema de predicción NMDS en R

4.2.1. Principal Component Analysis PCA

1. Se leen los datos con la función read.xls y se le asigna a un objeto llamado data

```
library(gdata)
read.xls("/home/topor/Descargas/ppca.xls")
data=read.xls("/home/topor/Descargas/ppca.xls")
```

2. Se eliminan las filas con NA

```
data=data[1:8]
str(data)
```

3. Se mira la correlación entre variables con la función cor() como se observa en la Figura 14.

```
cor(data[,1:8])
```

```
                edad         pam Proteinuria         ldh         ast
edad         1.000000000  0.08490485  0.35086216  0.4416447 -0.001078331
pam          0.084904850  1.00000000  0.28643601  0.3222493 -0.198112048
Proteinuria  0.350862164  0.28643601  1.00000000  0.5962791 -0.086765348
ldh          0.441644711  0.32224934  0.59627912  1.0000000 -0.126749731
ast         -0.001078331 -0.19811205 -0.08676535 -0.1267497  1.000000000
creatinina   0.332785012  0.11904879  0.56724138  0.4210237 -0.052287927
plaquetas   -0.048160307 -0.23446893 -0.08358099 -0.1000539 -0.133090816
eg           0.125165161  0.23243593  0.12881324  0.1494961 -0.047834023
             creatinina   plaquetas          eg
edad         0.33278501 -0.04816031  0.12516516
pam          0.11904879 -0.23446893  0.23243593
Proteinuria  0.56724138 -0.08358099  0.12881324
ldh          0.42102375 -0.10005394  0.14949608
ast         -0.05228793 -0.13309082 -0.04783402
creatinina   1.00000000 -0.09937160  0.09040401
plaquetas   -0.09937160  1.00000000  0.11114020
eg           0.09040401  0.11114020  1.00000000
```

Figura 14. *Correlación entre variables.*
Fuente: *Elaboración del autor.*

4. Se ajusta un PCA con la función prcomp() como se observa en la Figura 15.

```
pca1=prcomp(data[,1:8])
pca1
```

```
Standard deviations:
[1] 176.3690227  84.6914189  72.1734634  17.6355777  11.3601806   6.7683297
[7]   2.3997643   0.3075987

Rotation:
                PC1           PC2           PC3           PC4
edad        -0.017236187  0.0025455650  2.349211e-02 -0.0351491231
pam         -0.023500073  0.0324717516  1.832285e-02  0.2380928704
Proteinuria -0.913495224 -0.1127120827 -3.907208e-01 -0.0043194852
ldh         -0.401450791  0.1117295657  9.077617e-01 -0.0301178651
ast          0.009945955  0.0290904487 -2.639906e-02 -0.9698174407
creatinina  -0.001253191  0.0001295767 -2.909451e-05 -0.0003166462
plaquetas    0.058394195 -0.9863533244  1.473494e-01 -0.0237954309
eg          -0.002005931 -0.0035276686  3.112888e-03  0.0061959237
                PC5           PC6           PC7           PC8
edad        -0.059174305  0.996611832 -0.0336781322 -0.0061754175
pam          0.966725590  0.063096231 -0.0531337506  0.0025874009
Proteinuria -0.009502447 -0.007032687 -0.0002487547 -0.0011256786
ldh         -0.021321478 -0.030984492 -0.0008256525 -0.0003677505
ast          0.239595902 -0.019433394 -0.0052845121  0.0005159091
creatinina  -0.002819130  0.006127586  0.0038837052  0.9999688663
plaquetas    0.037227351  0.001238972 -0.0056258357  0.0003169556
eg           0.050942264  0.036844077  0.9979815555 -0.0039581460
```

Figura 15. *Dataset en Función del PCA*
Fuente: *Elaboración del autor.*

5. Se imprime el resumen del modelo

```
summary(pca1)
```

```
Importance of components:
                            PC1     PC2     PC3     PC4     PC5     PC6
Standard deviation     176.3690 84.6914 72.1735 17.63558 11.36018 6.76833
Proportion of Variance   0.7073  0.1631  0.1184  0.00707  0.00293 0.00104
Cumulative Proportion    0.7073  0.8704  0.9888  0.99589  0.99883 0.99987
                           PC7    PC8
Standard deviation     2.39976 0.3076
Proportion of Variance 0.00013 0.0000
Cumulative Proportion  1.00000 1.0000
```

Figura 16. *Resumen del modelo.*
Fuente: *Elaboración del autor.*

6. Se dibujan los dos primeros factores del PCA como se observa en la Figura 17.

```
biplot(pca1)
```

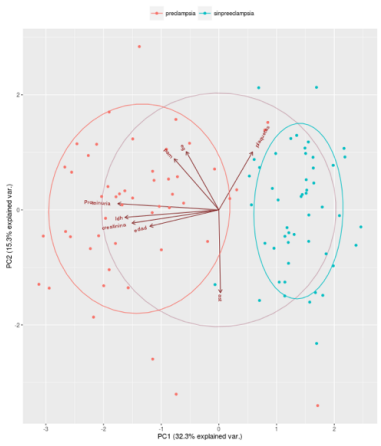


Figura 17. *Separación Factores predictores por clase.*
Fuente: *Elaboración del autor.*

Es posible notar de la anterior figura la separación de los factores, dentro de los cuales, los componentes principales (ldh, proteinuria, edad y creatinina) se alejan en el mismo sentido de las demás variables del modelo de datos.

4.2.2. Principal Coordinate Analysis PCoA o Non Metric Dimensional Scaling

1. Se utiliza la librería vegan y normalizan los datos
   library(vegan)
   dta=dat[1:8]
   dta
2. Se ingresa el dataset en la función metaMDS para hallar el escalamiento multi dimensional.
   NMDS = metaMDS(dta)
3. Se muestra la gráfica de las distancias entre factores almacenadas en la variable NMDS como se observa en la Figura 18.
   ordiplot(NMDS, type = "t")

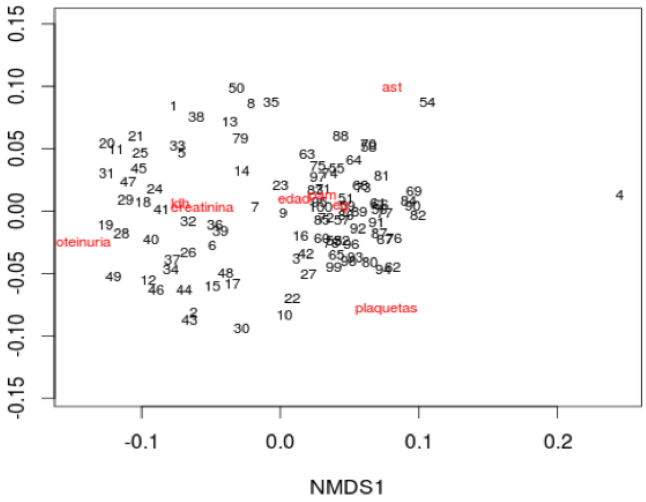


Figura 18. *Distancia entre Factores*
Fuente: *Elaboración del autor.*

4. Se selecciona un conjunto objetos en 2 clases por medio de la bondad de ajuste y se gráfica cómo se observa en la Figura 19 y Figura 20.
   par (mfrow = c(1,2))
   stressplot (NMDS)
   plot (NMDS, display = 'sites', type = 't', main = 'Goodness of fit')

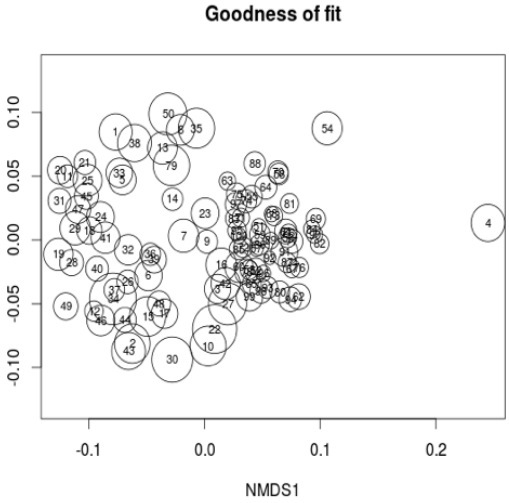


Figura 19. *Distancia entre las especies de un Dataset.*
Fuente: *Elaboración del autor.*

De la anterior Figura 19 se muestra la gráfica de NMDS2 vs NMDS1 en donde es posible notar las proximidades existentes entre un conjunto de objetos similarmente parecidos en distancia y que a la vez muestra interdependencia con otro conjunto de datos que serían representados como sin preeclampsia.

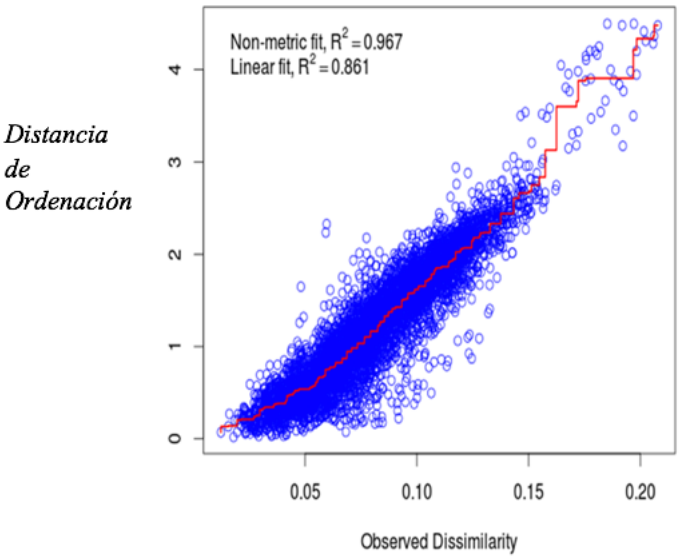


Figura 20. *Medida de ajuste para puntos de escalamiento multidimensional entre Distancia de Ordenación vs Disimilitud de Factores*
Fuente: *Elaboración del autor.*

En la Figura 20 se muestra una la gráfica de Distancia de Ordenación vs Disimilitud de Factores en donde la función stressplot es una envoltura para la función Shepard en el paquete MASS y traza las distancias de coordinación contra disimilitudes originales de los factores, y dibuja una línea de paso del ajuste no lineal. Además, se añade a la gráfica dos estadísticas de correlación que se extienden sobre la bondad del ajuste.

Resultados

Una vez que cada registro de pacientes enfermos y sanos ha sido procesado, se extraen los valores y se construye un árbol de decisión como grupo de entrenamiento usando el concepto de entropía de información. Los datos de entrenamientos son un grupo de ejemplos que tiene un conjunto de atributos o características del ejemplo representando así la clase a la que pertenece cada muestra.

En cada nodo el árbol C5.0 elige un atributo de los datos que más eficazmente dividen el conjunto de muestras en subconjuntos enriquecidos en una clase u otra. Este criterio es formalizado para la ganancia de información según la diferencia de entropía que resulta en la elección de un atributo para dividir los datos. El atributo con mayor ganancia de información normalizada se elige como parámetro de decisión.

En total se escogió un conjunto de 100 datos simulados bajo restricciones de registros reales de pacientes, de los cuales 80 correspondiente al 75% de la muestra pertenecieron al conjunto de entrenamiento en donde se generaron las reglas de clasificación que constituyen la base para encontrar el factor predictor de la preeclampsia y los otros 20 corresponden al 25% de la muestra que sirvió como validación.

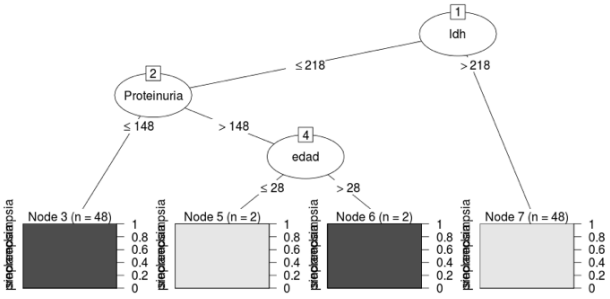


Figura 21. *Árbol C5.0 - 3 niveles*
Fuente: *Elaboración del autor.*

Al granular el árbol C5.0 se encuentra que el segundo factor más importante es la proteinuria seguido de la edad. Estableciendo los valores de la variable en la clasificación Proteinuria <= 150 y > 150 y los valores de Edad <= 28 y > 28 de (sano) sin preeclampsia y (enfermo) preeclampsia mostrada en la Figura 12. Árbol C5.0 - 3 niveles .

```
Console ~/

Call:
C5.0.default(x = datar[1:80, -10], y = datar[1:80, 10])


C5.0 [Release 2.07 GPL Edition]        Wed Feb 10 13:08:21 2016
-------------------------------

Class specified by attribute `outcome'

Read 80 cases (10 attributes) from undefined.data

Decision tree:

ldh <= 218: sinpreeclampsia (44/2)
ldh > 218: preclampsia (36)


Evaluation on training data (80 cases):

          Decision Tree
        ----------------
        Size      Errors

           2    2( 2.5%)   <<


         (a)   (b)    <-classified as
        ----  ----
          36     2    (a): class preclampsia
                42    (b): class sinpreeclampsia


       Attribute usage:

       100.00% ldh


Time: 0.0 secs
```

Figura 22. *Resultados del Entrenamiento C5.0*
Fuente: *Elaboración del autor.*

Para la cual la variable con más ganancia de información fue ldh (lactato deshidrogenasa) enzima catalizadora encontrada en muchos tejidos del cuerpo. Estableciendo los valores de la variable en la clasificación ldh <= 218 y > 218 de (sano) sin preeclampsia y (enfermo) preeclampsia mostrada en la Figura 13. Resultados del Entrenamiento C5.0.

Según los resultados obtenidos con el algoritmo NMDS están subdivididos en 2 parte PCA Y PCoA los cuales realizan separación de los datos. Utilizando la técnica PCA propuesta conceptualmente por Cailliez (1983), Legendre (1998) y Lingoes (1971) e implementada por Gower (2005), es posible observar dentro de un dataset que componentes son

denominados principales como se muestra en la Figura 17. Separación Factores predictores. Donde se muestran que las variables que más impacto tienen en dataset son (ldh, proteinuria, plaquetas) y las cuales se alejan de las demás variables del modelo de datos.

Utilizando la técnica PCoA y recibiendo la matriz de distancias del PCA fue posible identificar qué factores pertenecían más a cada especie (preeclampsia, sin preeclampsia). Se observó que los factores como edad, pam (presión arterial media), eg (edad gestacional) no se separaban del conjunto de datos o especie sin preeclampsia como se muestra en la Figura 9. Distancia entre Factores. Además realizando una Bondad de ajuste, es posible notar las proximidades existentes entre un conjunto de objetos similarmente parecidos en distancia y que a la vez muestra interdependencia con otro conjunto de datos son representados como sin preeclampsia por los cual se logró separar las dos especies o clases como se muestra en la Figura 10. Distancia entre las especies de un dataset. Luego se utiliza la función stressplot de envoltura para la función Shepard en el paquete MASS. En esta se traza las distancias de coordinación contra disimilitudes originales, y dibuja una línea de paso del ajuste no lineal como se muestra en la siguiente Figura 11. Medida de ajuste para puntos de escalamiento multidimensional.

Conclusiones

El sistema de clasificación de preeclampsia en mujeres embarazadas desarrollado muestra cómo pueden implementar técnicas de inteligencia artificial en medicina enfocadas a las áreas ginecología, para el reconocimiento automático de enfermedades en las entidades de atención médicas de tercer nivel.

A partir de los resultados obtenidos en la implementación de este proyecto para el reconocimiento de la preeclampsia se puede concluir lo siguiente.

La etapa de adquisición de la información juega un papel importante en el proceso de predicción de la preeclampsia. Puesto que depende en gran medida del extracto social al que pertenece la entidad médica a donde son atendidas estas mujeres lactancia.

- Es importante tener en cuenta, que se deben eliminar las variables que no aportan ganancia de información dentro del árbol entrenado con el algoritmo C5.0.
- La variable LDH resulto ser el factor más adecuado para la predicción de preeclampsia. Esto se puede reafirmar visualmente puesto que es la variable con mayor ganancia de información de las elegidas como predictoras (ver Figura 9. Árbol entrenado C5.0).
- Tanto en el algoritmo C5.0 como en las técnicas PCA y NMDS se visualizó que los factores LDH, Proteinuria y Edad son los factores más adecuados para la predicción de preeclampsia.
- La clasificación de preeclampsia aplicando el algoritmo de árboles de decisión C5.0 presento un porcentaje de error del 2.5% dando un porcentaje de acierto del 97.5%. Una vez el árbol de decisión ha sido entrenado, el reconocimiento de la enfermedad es automático. Además, según las pruebas de cribaje explicado por el autor (Morgado, 2008), utilizando las técnicas tradicionales (Mancuernas con presión arterial media y MAP con Factores de riesgo) para la predicción de preeclampsia, se encuentra una tasa de detección del 12% y 62.5% respectivamente.
- Comparando el porcentaje de acierto del modelo explicado en este documento, con el presentado por el autor (Neocleous), que aplicando la técnica de minería datos, redes neuronales obtuvo un porcentaje de acierto del 83,6% de la preeclampsia y en la prueba de ajuste del 93,6%, por los cual se observa una mayor eficiencia en el resultado del modelo propuesto en esta investigación.

Referencias Bibliográficas

Bautista, A. (2012). Hipertensión arterial asociada con el embarazo.

Cailliez, F. (1983). The analytical solution of the additive constant problem. Springer, 48(2).

Cifuentes, R. (2006). Hipertensión arterial y embarazo En:Obstetricia de alto riesgo. Distribuna, 447-484.

Cifuentes, R. (2009). Preeclampsia-eclampsia. Diagntico y manejo En: Ginecologia y Obstetricia Basadas en las Nuevas Evidencias. Distribuna, 345-348.

Cifuentes, R. (2010). Preeclampsia-eclampsia diagnòstico y manejo. Bogota: Altavoz.

Cook, E. & Goldman, L. (1984). Empiric comparison of multivariate analytic techniques: advantages and disadvantages of recursive partitioning analysis. Journal of chronic diseases, 37(9), 721-731.

Dustin, T. & Dunsmuir, B. A. (2014). Development of mHealth Applications for Pre-Eclampsia Triage.

Etchegaray Adolfo, S. M. (2012). Predicción de preeclampsia en el primer trimestre- validación prospectiva preliminar de un método de screening combinado.

García, D., Porto, R., Simancas, R., Molina, M. & Hernández, H. (2016). Sistemas de pronosticos agrícolas, una apuesta a la optimizacion de sus insumos. Caso del tuberculo de la malanga. Produccion+ Limpia, 11(1), 141-149.

Garovic, V. D. (2014). The Role of Angiogenic Factors in the Prediction and Diagnosis of Preeclampsia Superimposed on Chronic Hypertension. Editorial Commentary, 740-746.

Gower, J. C. (2005). Principal Coordinates Analysis Encyclopedia of Biostatistics. Encyclopedia of biostatistics.

Harskamp, Z. G. (2007). Preeclampsia: At Risk for Remote Cardiovascular Disease. Am J Med Sci, 334:291-295.

Legendre, P. (1998). Numerical Ecology (1 ed.). Elsevier.

Lingoes, J. C. (1971). Some boundary conditions for a monotone analysis of symmetric matrices. Ann Arbor: Springer.

Loaiza, M. F., Solano, R. P., Simancas, R., & Ojito, V. H. (2017, April). Modeling Demand for Air Cargo in the Colombian Context. In 2017 International Conference on Advanced Materials Science and Civil Engineering (AMSCE 2017). Atlantis Press.Morgado, R. (2008). Metodos de predicción en la preeclampsia. DOSSIER.

Neocleous, C. K. (2009). Neural networks to estimate the risk for preeclampsia occurrence. IEEE, 7.

Porto Solano, R., Porto Barceló, R., Corredor, A., Cortez, J., Echeverri, C., De los Rio, J., & Herrera, E. (2017). Framework ágil para el control de recetas médicas que utiliza la tecnología NFC (FARM). Revista LASALLISTA de Investigación, 206-216.

Quinlan, J. (2014). C4. 5: programs for machine learning. San Francisco: Elsevier.

S.Y. Leemaqz, G. D. (2013). Tiered Prediction System for Preeclampsia: an integrative application of multiple models.

Sibai B, D. G. (2005). Pre-eclampsia. Lancet, 785-99.

Vázquez, C. (21 de octubre de 2011). EROSKI CONSUMER Obtenido de: http://www.consumer.es/web/es/bebe/embarazo/sintomas/2011/10/20/204305.php.

Verlohren, S. (2012). Angiogenic growth factors in the diagnosis and prediction of pre-eclampsia.