

Reseñas de algunos trabajos de pregrado de las carreras de Matemáticas y Ciencias de la Computación Universidad Nacional de Colombia Sede Bogotá. 2023-I

Dirección de Área Curricular^a

Carrera de Matemáticas

1. Estudio de operadores sobre $l^2(\mathbb{Z})$ y otros espacios de Dimensión Infinita (Teorema Espectral)

Estudiante **Sergio Arango Motta***

Director *Guillermo Rodríguez Blanco***

Emails *sarangomo@unal.edu.co, **grodriguez@unal.edu.co

RESUMEN. Throughout the years, Linear Algebra has incurred in various generalizations and abstractions in its core concepts. These generalizations allow for vector spaces to be infinitely dimensional. What once was a simple matrix denoting a linear transformation, can be generalized to be an operator. This work will bring concepts from Operator Theory and Spectral Theory to show how concepts such as eigenvectors and eigenvalues generalize. Various Banach and Hilbert spaces will be considered. We will draw the similarities the Spectral Theorem has in the finite and infinite cases.

2. Pasantía en el Fondo Latinoamericano de Reservas

Estudiante **Ana Paulina Castillo Velasquez***

Director *Francisco Albeiro Gómez Jaramillo***

Emails *ancastillov@unal.edu.co, **fagomezj@unal.edu.co

RESUMEN. El Fondo Latinoamericano de Reservas (FLAR) es una organización financiera internacional que fue creada en 1978 por los bancos centrales de América Latina y el Caribe. Su objetivo principal es proporcionar un mecanismo de cooperación financiera regional a través de

¹Departamento de Matemáticas, Universidad Nacional de Colombia, Bogotá, Colombia

^acoocurmat_fcbo@unal.edu.co

préstamos de corto plazo y asesoramiento a los bancos centrales para ayudar a sus miembros a enfrentar desafíos económicos.

Para esto es necesario orientar planes de acción conforme las proyecciones económicas del comportamiento del mercado; conllevando a un proceso de análisis de datos, donde es imperante la precisión y agilidad. Por lo anterior es pertinente la implementación de medidas de Inteligencia Empresarial (IE), como extracción, transformación y carga (ETC) enfocadas a la economía desde una perspectiva matemática, con el fin de crear, mejorar e implementar algoritmos dentro del tratamiento de la información para su eficiente lectura.

Conforme lo anterior, la tarea a realizar se puntualiza en la automatización del tratamiento de datos haciendo una traducción desde el método clásico utilizado en la empresa a un lenguaje de programación como lo es *Python*, pasando por una verificación de las fórmulas utilizadas en los cálculos. Particularmente abordando la proyección del valor de la tasa de inversión ofrecida por el Tesoro de los E.E.U.U.

3. Algunas propiedades de las funciones enteras de tipo exponencial

Estudiante **Carlos Alberto Garzón Gutiérrez***

Director *Germán Preciado López***

Emails * cgarzongu@unal.edu.co, ** gpreciadol@unal.edu.co

RESUMEN. El teorema de Paley-Wiener nos proporciona condiciones necesarias y suficientes para que una función en $L^2(\mathbb{R})$ tenga transformada de Fourier con soporte compacto. Las funciones que cumplen dichas propiedades se conocen como funciones enteras de tipo exponencial. En este trabajo vamos a enunciar y probar el Teorema de Paley-Wiener, lo cual nos permitirá obtener el Teorema de Muestreo de Shannon-Whitaker. Adicionalmente, estudiaremos otras propiedades que cumplen las funciones enteras de tipo exponencial.

Palabras clave. Teorema de Paley-Wiener, funciones de tipo exponencial, transformada de Fourier, Teorema de Shannon-Whitaker, Principio de Phragmén-Lindelöf.

4. Sobre el teorema de Riemann-Roch para curvas algebraicas

Estudiante **Daniel Steven Martínez Lizarazo***

Director *Armando Reyes***

Emails * dmartinezli@unal.edu.co, ** mareyesv@unal.edu.co

RESUMEN. En este trabajo estudiamos el teorema de Riemann-Roch para curvas algebraicas después de hacer una introducción a varios fundamentos elementales que se encuentran en la geometría algebraica. Además, se mencionan los conceptos de divisores en curvas algebraicas proyectivas

no singulares e irreducibles, el género de dicha curva y las formas diferenciales en la misma. Con el teorema de Riemann-Roch se demuestra un proceso inductivo, y luego, hay propuestas de posibles extensiones y aplicaciones futuras del teorema en diferentes áreas de las matemáticas.

5. On the noncommutative projective geometry of some quantized algebras

Estudiante **José Sebastián Naranjo Spadafora***

Director *Armando Reyes***

Emails *jonaranjos@unal.edu.co, **mareyesv@unal.edu.co

RESUMEN. In this thesis we present a general study of projective space, beginning with a look at the essence of its construction, then presenting relevant results of its algebraic geometry and briefly discussing its purely geometric aspect. Later, inspired by the commutative case, we will obtain a correspondence between a special type of modules in some non-commutative algebras.

6. Ecuaciones diferenciales parciales que satisfacen polinomios ortogonales en dos variables

Estudiante **María Alejandra Pérez Sánchez***

Director *Herbert Alonso Dueñas Ruíz***

Emails *maperezsa@unal.edu.co, **haduenasr@unal.edu.co

RESUMEN. En este documento se estudian los polinomios ortogonales clásicos en una y dos variables [2, 3], los cuales satisfacen ciertas ecuaciones diferenciales ordinarias y parciales respectivamente. Se estudia la conexión mostrada en [1] entre los polinomios ortogonales clásicos en una variable y su clasificación, y un tipo de polinomios en dos variables definido a partir de productos de los de una variable, estudiando así la forma de las ecuaciones diferenciales parciales (EDP) que éstos satisfacen a partir de las ecuaciones diferenciales ordinarias (EDO) de los polinomios univariados. Se presentan diferentes ejemplos de polinomios ortogonales clásicos en una y dos variables generados en Matlab, y las ecuaciones diferenciales que satisfacen.

Palabras clave. Polinomios ortogonales, polinomios ortogonales bivariables, polinomios ortogonales clásicos, producto interno, ecuaciones diferenciales, ecuaciones diferenciales parciales.

- [1] J. C. García-Ardila and M. E. Marriaga, *Sobolev orthogonality of polynomial solutions of second-order partial differential equations.*, Computational and Applied Mathematics, Springer, (2023), **42**, pages 13, <https://doi.org/10.1007/s40314-022-02152-2>.
- [2] T. Chihara, *An introduction to orthogonal polynomials*, Courier Corporation, 1978.

- [3] Y. Xu and C. Dunkl, *Orthogonal polynomials of several variables*, Cambridge University Press, 2014, DOI: 10.1016/j.aam.2004.03.002.

7. Pasantía en el área de testing de la empresa Endava PLC

Estudiante **Luz Mariana Pulido Arcia***

Director *Francisco Albeiro Gómez Jaramillo***

Emails *lpulidoa@unal.edu.co, **fagomezj@unal.edu.co

RESUMEN. El testing de software es una parte crucial en el desarrollo y mantenimiento de una aplicación, ya que permite detectar y corregir errores y defectos antes de que lleguen a los usuarios finales. El testing de regresión en particular, es una técnica de testing que se enfoca en asegurar que los cambios realizados en el software no afecten la funcionalidad existente. Aunque es una práctica importante, puede ser costosa en términos de tiempo y recursos. Es por lo anterior que resulta fundamental optimizar el proceso de testing de regresión, ya sea disminuyendo los casos de prueba a ejecutar o priorizando algunos en el orden de ejecución. Se presenta una aproximación de solución para este problema mediante su planteamiento como problema de optimización multi-objetivo considerando como objetivo la métrica *cobertura de impacto*.

8. On lattices and functional Alexandroff topologies

Estudiante **Jhonatan Camilo Rodríguez Porras***

Directora *Ibeth Marcela Rubio Perilla***

Emails *jrodriguezpo@unal.edu.co, **imrubio@unal.edu.co

RESUMEN. En este trabajo de grado se presentan algunas propiedades de los espacios funcionales de Alexandroff (o espacios primarios), estudiados e introducidos de manera independiente, pero casi simultánea hacia 2010 por Ayatollah, Golestani y Echi, y definidos a partir de una función de un conjunto en sí mismo. También se presentan unos resultados del estudio de la colección de estos espacios definidos sobre un conjunto dado, desde la perspectiva de los retículos. Además, para el caso finito se usaron algunas herramientas computacionales para implementar varios de los resultados presentados.

9. Pasantía como Data Analyst en Siigo

Estudiante **Edward Daniel Soto Mejía***

Director *Francisco Albeiro Gómez Jaramillo***

Emails *edsotom@unal.edu.co, **fagomezj@unal.edu.co

RESUMEN. Entender los hábitos de los clientes puede convertirse en una oportunidad de negocio para mejorar la calidad del servicio y la experiencia del cliente. En particular, esta oportunidad se refleja en comprender cómo los clientes utilizan los productos que una empresa ofrece. Por esta

razón, el equipo de Data Analytics de Siigo se propuso la tarea de plantear un modelo estadístico/matemático capaz de describir el comportamiento de uso que los clientes dan a los productos.

El modelo planteado fue el clásico modelo de aprendizaje no supervisado K-Means, cuyo desempeño se evaluó mediante el Silhouette Score. Para obtener los resultados del modelo, se iteró en un proceso de análisis de datos para depurar las bases de datos de Siigo, y se aplicaron técnicas de ingeniería de características para evitar un posible sesgo del modelo. Se utilizó una normalización de tipo Min y Max junto con un análisis de componentes principales (ACP). El flujo de datos varió desde el uso de herramientas como DataBricks de Microsoft para depurar los datos, hasta scikit-learn para implementar la ingeniería de características y construcción del modelo.

Los resultados muestran un amplio uso por parte de los clientes de Siigo hacia sus productos. Se concluye que los clientes están satisfechos con los productos que Siigo ofrece. Sin embargo, se plantea la posibilidad de un posible colapso de la plataforma debido al alto flujo de datos que generan los usuarios.

10. **Las matemáticas detrás de algunos aspectos de economía en nuestra cotidianidad**

Estudiante **Laura Natalia Suarez Montilla***

Director *Jorge Mauricio Ruiz Vera***

Emails *lsuarezmo@unal.edu.co, **jmruizv@unal.edu.co

RESUMEN. La Imprenta de Billetes tiene como función la producción del papel moneda, para lo que requiere tener abastecimiento de los insumos necesarios. Con el fin de dar un manejo adecuado a los inventarios de tales insumos se buscaba implementar la metodología de Planeación de Requerimientos de Materiales Bajo Demanda (DDMRP, por sus siglas en inglés), en la que destacan el consumo promedio diario y la demanda futura; por lo que se realizaron estimaciones de los consumos de diferentes insumos mediante Mínimos Cuadrados Ordinarios, Modelo Lineal Robusto y Random Forest. Para varios de los insumos se pudo validar el modelo usado comparando los consumos diarios pronosticados durante dos meses. Tras dicha validación y una vez realizada la verificación con los responsables del área, se incluyen las estimaciones obtenidas en la metodología DDMRP, la cual se empieza a implementar para los insumos de la última etapa de producción. Finalmente se emplea Power BI para visualizar de manera más clara estos resultados.

11. **Littlestone in Algebraically Closed Fields**

Estudiante **José Alfredo Urrea Moreno***

Director *Andrés Villaveces Niño***

Emails *jourream@unal.edu.co, **avillavecesn@unal.edu.co

RESUMEN. Cuando se estudia la dimensión de Littlestone en conjuntos definibles dados por $\varphi(x; b)$, la finitud de esta dimensión es equivalente a que dicha fórmula sea estable. Esta tesis se enfoca en el caso particular de fórmulas positivas en Campos Algebraicamente Cerrados (ACF). En el caso bidimensional, la dimensión de Littlestone puede ser acotada por $4 \log_2(dm) + \log_2(m) + 1$, donde d representa el grado polinomial y m es el número de componentes irreducibles. El objetivo es desarrollar métodos para el estudio de la teoría de aprendizaje en línea en teorías no estables. Para lograr esto, se examina la eliminación de cuantificadores en ACF y se proponen posibles desarrollos para el caso general utilizando Bases de Gröbner y órdenes monomiales.

12. Geometría diferencial básica de las curvas modulares

Estudiante **Heldert Villegas***

Director *Leonardo Cano***

Emails *hvillegas@unal.edu.co, **lcano@unal.edu.co

RESUMEN. Una curva elíptica está dada por una ecuación de la forma $y^2 = 4x^3 - ax - b$. Estas son muy importantes en diversas áreas de las matemáticas como la criptografía y la teoría de números. Existe una equivalencia entre la categoría de curvas elípticas y rejillas sobre los números complejos, estas son conjuntos de la forma $\omega_1\mathbb{Z} \oplus \omega_2\mathbb{Z}$ donde ω_1 y ω_2 son números complejos linealmente independientes bajo \mathbb{R} . Sin embargo una rejilla puede tener distintos generadores, en particular toda rejilla es equivalente a una de la forma $\tau\mathbb{Z} \oplus \mathbb{Z}$ donde $\tau \in \mathbb{H} := \tau \in \mathbb{C} : \text{Im}(\tau) > 0$. Dos elementos $\tau, \tau' \in \mathbb{H}$ generan la misma rejilla si $\tau' = \frac{a\tau+b}{c\tau+d}$ donde $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z})$.

Las curvas modulares son cocientes de la forma \mathbb{H}/Γ (o H^*/Γ en su forma compactificada), donde Γ es un subgrupo de congruencia de $SL_2(\mathbb{Z})$, es decir, que contiene un grupo de la forma

$$\Gamma(N) := \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z}) : \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \pmod{N} \right\}$$

por tanto a cada punto de las curvas modulares corresponde una curva elíptica, haciendo de estas el espacio moduli para las curvas elípticas. Estas son superficies de Riemann, y por tanto pueden ser estudiadas haciendo uso de las técnicas y teoremas de la geometría diferencial.

En este trabajo se calculó el género de las curvas modulares haciendo uso del teorema de Riemann-Hurwitz, del cual presentamos una prueba usando el teorema de Gauss-Bonnet. Esto se desarrolló haciendo énfasis especial en los detalles con el fin de obtener una presentación sencilla de leer para un estudiante avanzado de pregrado. Al final se obtuvo la

siguiente formula que relaciona la información aritmetica de N con la información topológica de $\mathbb{H}^*/SL_2(\mathbb{Z})$ mediante la formula

$$g_N = 1 + \frac{N^2(N-6)}{24} \prod_{p|N} \left(1 - \frac{1}{p^2}\right).$$

13. On the optimization of deep networks

Estudiante **Jaime Andres Zamora Avilez***

Director *Francisco Albeiro Gómez Jaramillo***

Emails *jzamora@unal.edu.co, **fagomezj@unal.edu.co

RESUMEN. En este trabajo de revisión de un artículo de Tomaso Poggio se estudia el fenómeno de regularización implícita que existe en redes neuronales profundas para clasificación binaria con función de activación ReLU y función de pérdida de tipo exponencial. Definimos el problema en cuestión, exponiendo su motivación y algunos prerrequisitos técnicos necesarios para poder comprenderlo. Posteriormente pasamos a estudiar en detalle algunos de los resultados principales enunciados en el paper que motivó este trabajo, y estudiamos sus implicaciones en el campo de redes neuronales profundas. Finalmente, realizamos algunos experimentos con PyTorch para redes neuronales sobreparametrizadas con y sin regularización, para tareas de clasificación, con el fin de estudiar y contrastar los resultados estudiados en el trabajo. Concluimos que existe una regularización implícita para las redes neuronales estudiadas bajo las suposiciones hechas. Se requiere un estudio que permita generalizar algunos de los resultados estudiados en el trabajo. A pesar de existir un control implícito de la complejidad, se observó que introducir una regularización explícita permite acelerar la convergencia del gradiente a soluciones óptimas.

Palabras clave. Redes neuronales, Sobreparametrización, Regularización, Sistemas dinámicos.

Carrera de Ciencias de la Computación

1. SDF-based uncertainty representation for directional information in dMRI

Estudiante **Tania Valentina Castillo Delgado***

Director *Javier Ricardo Guaje Guerra***, Codirector *Francisco Albeiro Gómez Jaramillo****

Emails **tvcastillod@unal.edu.co*, ***jrguajeg@gmail.com*,

****fagomezj@unal.edu.co*

RESUMEN. Diffusion Magnetic Resonance Imaging (dMRI) is a non-invasive imaging technique used by physicians and scientists to measure the diffusion of water molecules in biological tissue. This technique allows modeling axonal directional information at a millimetric level through Diffusion Tensor Imaging (DTI). This method combines the information of multiple dMRI volumes into a tensor matrix, commonly visualized as ellipsoids. Traditionally, these ellipsoids are rendered as polygonal constructs composed of many vertices and triangles, which at a large scale, might severely impact the interactivity of such visualizations. In this work, we propose a novel way to render these tensor ellipsoids and their associated uncertainty using state-of-the-art computer graphics techniques, providing close to real-time performance, as shown in our results.

2. Consultor Aplicaciones en IBM

Estudiante **Jorge Luis Castillo Orduz ***

Director *Agustin Moreno Cañadas***

Emails **jlcastilloo@unal.edu.co*, ***amorenoca@unal.edu.co*

RESUMEN. Este documento presenta un resumen completo de mi experiencia de pasantía de pregrado en IBM, donde trabajé como Consultor de Aplicaciones. Se exploran los conocimientos adquiridos, las tareas realizadas y las aplicaciones prácticas a lo largo de la pasantía.

El trabajo se llevó a cabo en el contexto de la transformación tecnológica de Danske Bank, el banco más grande de Dinamarca y los países nórdicos. El proyecto se centró en desarrollar una solución de nube híbrida que permitiera al banco mantener su sistema actual mientras actualiza sus procesos.

Los objetivos planteados por IBM para la pasantía incluyeron comprender el proceso de desarrollo y gestión de servicios tecnológicos, construir y configurar servicios de aplicaciones utilizando arquitectura limpia y metodologías ágiles, e implementar buenas prácticas de optimización de recursos tecnológicos. Estos objetivos fueron cumplidos durante el transcurso de la pasantía.

El proceso de la pasantía se dividió en cuatro partes principales: conocer IBM, capacitación, retos prácticos y proyecto. Durante la etapa de capacitación, se completaron varios cursos que abarcaron temas como Agile, Kubernetes, Docker, Git, seguridad cibernética y arquitectura de microservicios.

La pasantía en IBM me brindó la oportunidad de aprender sobre la historia y los planes de la empresa, así como de adquirir conocimientos y habilidades en tecnologías y prácticas relevantes para la industria. Además de cumplir con los objetivos establecidos, mi objetivo personal fue ser contratado y trabajar a tiempo completo con IBM debido a las oportunidades de crecimiento profesional que ofrece la compañía.

En resumen, esta pasantía en IBM me permitió adquirir experiencia en consultoría de aplicaciones y desarrollar habilidades clave en tecnologías de vanguardia. Fue una experiencia enriquecedora que contribuyó significativamente a mi formación y desarrollo profesional.

3. Pasantia IBM

Estudiante **Christian Camilo Pabon Useche***

Director *Francisco Albeiro Gríez Jaramillo***

Emails *ccpabonu@unal.edu.co, **fagomezj@unal.edu.co

RESUMEN. La pasantía que se está presentando para Ciencias de la Computación es realizada en la empresa IBM COLOMBIA, empresa dedicada a prestar servicios de tecnología. El trabajo consiste en mostrar lo que hice en IBM y la presentación de un resto que abarca los conocimientos adquiridos en la pasantía.

Palabras clave. IBM Colombia, IBM MQ, Integración.

4. Inteligencia en metricas de soporte en la empresa Uber

Estudiante **Gabriel Alejandro Quinche Velandia***

Director *Francisco Albeiro Gómez Jaramillo***

Emails *gquinche@unal.edu.co, **fagomezj@unal.edu.co

RESUMEN. En este trabajo de grado, presentaré los conocimientos y experiencias adquiridos durante mi pasantía en Uber Colombia, donde me enfoqué en analíticas para el soporte al cliente y la reducción de costos. Uno de los principales desafíos que encontré fue la necesidad de automatizar y simplificar el monitoreo del alto volumen de llamadas de servicio al cliente. Para abordar este desafío, desarrollé una solución práctica que consistió en un panel de datos o dashboard que presentaba los cambios en la métrica de tasa de defectos con una vista simplificada de los aumentos mensuales y semanales para un análisis y reacción eficiente. Este proyecto tuvo un impacto positivo al reducir los contactos al soporte, el fraude y

umentar la eficiencia general. También se presentan dos secciones relacionadas con la alineación de precios mediante optimización convexa y la vectorización de palabras para análisis cualitativos. Estas dos secciones resaltan preguntas ad hoc que, aunque no se realizaban semanalmente, estaban relacionadas con procesos al menos anuales, y su automatización permitió ahorrar tiempo en los equipos de la región y una mejor comprensión de los problemas.

Palabras clave. Customer support, Automation, Data analysis, Cost reduction, Optimization, Data visualization.

5. Wed Developer Intern, Pasantía en ZAGALABS SAS

Estudiante **Daniel Ricardo Rodríguez Olarte***

Director *Francisco Albeiro Gómez Jaramillo***

Emails *drodriguezol@unal.edu.co, **fagomezj@unal.edu.co

RESUMEN. Este trabajo de grado presenta el desarrollo de una arquitectura de procesamiento de datos segura y escalable, basada en la pasantía realizada en ZagaLabs. El proyecto EDS (ERP Data Shuttle) tiene como objetivo principal diseñar y desarrollar una solución eficiente para el traspaso y procesamiento de datos de clientes reales de la empresa Second Nature. Durante esta pasantía, se han enfrentado diversos desafíos técnicos y de liderazgo, lo que ha permitido adquirir habilidades valiosas en el campo de la ciencia de la computación.

Problemática

La empresa Second Nature comenzó originalmente como un proveedor de filtros de aire. Con el objetivo de adquirir más clientes, ampliaron sus servicios ofreciendo el cambio de filtros a arrendatarios. Con el tiempo, esta iniciativa evolucionó y se creó RBP (Resident Benefits Program), que es un paquete de beneficios para residentes.

Actualmente, los responsables de Second Nature desean crear RXP (Resident Experience Platform), un sistema centralizado y automatizado que se encargue de todo el proceso de solicitud de contratos de arrendamiento (leases). Los property managers de otras empresas utilizarán este sistema para gestionar y enviar los contratos de arrendamiento a los residentes. El objetivo es que los residentes puedan ver, pagar, firmar, modificar y realizar consultas a los propietarios, entre otras acciones, todo a través de esta plataforma.

Sin embargo, no toda la información necesaria para RXP está disponible de inmediato. Parte de la información debe migrarse o sincronizarse desde RBP (el otro servicio de Second Nature), mientras que otra parte se obtiene de páginas de arrendatarios, sitios web gubernamentales y otras fuentes de datos abiertos.

Debido a la gran cantidad y diversidad de los datos que se deben extraer, se creó una aplicación independiente llamada EDS. Esta aplicación se encarga de proporcionar los datos a RXP de manera más fácil, transparente y limpia. EDS obtiene los datos de diversas fuentes en línea.

La limpieza de los datos es un proceso fundamental, ya que provienen de diversas fuentes y requieren un formato estandarizado. Para evitar sobrecargar a RXP, EDS se encarga de esta tarea. La interacción entre EDS y RXP es bidireccional: en su mayoría, EDS proporciona datos a RXP, pero en ciertos casos, para sincronizar información específica y manejar aspectos particulares de RXP, EDS necesita llamar a servicios de RXP por motivos de autenticación, seguridad o por cuestiones de tiempo.

El proyecto EDS aborda la problemática de contar con una arquitectura que permita leer y escribir grandes volúmenes de datos de manera eficiente y escalable. El objetivo es garantizar la integridad, precisión y actualización de los datos provenientes del ERP de Second Nature, optimizando los procesos de extracción, transformación y carga de información.

Materiales y métodos

Durante el desarrollo del proyecto EDS, se utilizaron diversos servicios de AWS (Amazon Web Services) para construir una arquitectura segura, escalable y eficiente.

El web scraping se convirtió en una herramienta esencial para la adquisición de datos necesarios en el proyecto. En esencia, el web scraping es una técnica utilizada para extraer información valiosa de sitios web de manera automatizada.

En nuestro caso, utilizamos el web scraping para recopilar datos relevantes de diversas fuentes en línea que estaban relacionadas con el dominio específico de nuestro proyecto. Para lograr esto, hicimos uso de librerías especializadas, las cuales nos permitieron analizar la estructura HTML de las páginas web y extraer los datos necesarios de forma automatizada.

Resultados

El proyecto EDS desarrolló una arquitectura que consta de los siguientes pasos:

- a) Se creó un servicio web utilizando un framework de Python para extraer y enviar la información requerida desde diversos sitios en línea.
- b) El servicio web se encapsuló en un contenedor de Docker para facilitar su despliegue y portabilidad.
- c) Posteriormente, se alojó el contenedor en un dominio específico con AWS (Amazon Web Services), permitiendo el acceso a través de la web.

- d) Se estableció una ejecución periódica del servicio web, lo que implica que se programó para que se ejecute de forma automatizada en intervalos regulares.
- e) El servicio web activa diferentes servicios de AWS (Amazon Web Services), los cuales se encargan de almacenar, organizar y transformar los datos de manera segura y eficiente. Estos servicios pueden incluir, por ejemplo, almacenamiento en la nube, bases de datos, servicios de procesamiento, etc.
- f) Finalmente, se entrega la información necesaria a RXP para construir el sistema deseado correspondiente.

<https://drive.google.com/file/d/1Lh1u6nY6j5h-7bqHMVOyGhgKbpsl0QcC/view> Link Esquema de arquitectura

En adición, hemos desarrollado servicios basados en APIs con Flask y Django, que se despliegan en AWS para actualizar nuestra base de datos de forma automatizada. Estos servicios funcionan como endpoints accesibles a través de una API y se ejecutan periódicamente para obtener los datos más recientes y realizar las actualizaciones necesarias en la base de datos. Esta solución nos permite mantener la integridad y precisión de la información en nuestro sistema, al mismo tiempo que automatizamos el proceso de actualización para mejorar la eficiencia y optimizar las operaciones.

Discusiones y conclusiones

- a) La arquitectura desarrollada demostró ser eficiente y escalable, permitiendo el procesamiento de grandes volúmenes de datos de manera segura y confiable. La utilización de servicios de AWS fue clave para lograr la integración y optimización de los procesos.
- b) El proyecto EDS ha alcanzado sus objetivos y metas propuestas de manera satisfactoria, cumpliendo con los objetivos propuestos y generando una solución sólida para el traspaso y procesamiento de datos de la empresa Second Nature.
- c) A pesar de la ausencia de formación específica en arquitectura de software, pude aportar al proyecto mediante el aprovechamiento de mis habilidades y conocimientos adquiridos en ciencias de la computación. Mi capacidad de aprendizaje rápido, adaptabilidad, resolución de problemas, colaboración, pensamiento crítico y creativo fueron fundamentales para contribuir con éxito al desarrollo de la arquitectura de procesamiento de datos segura y escalable en el proyecto EDS.
- d) Mis logros y contribuciones en el proyecto no pasaron desapercibidos por la empresa. Como resultado, me ofrecieron un empleo en Zaga-Labs, con un salario competitivo que reconoce mi aporte al proyecto y mis habilidades adquiridas en mi carrera.

- e) La arquitectura implementada demuestra un crecimiento continuo y su escalabilidad ofrece una base sólida para la evolución y futuros desarrollos del proyecto.

Referencias

- a) ZagaLabs. Sitio web: <https://www.zagalabs.com/>
- b) SecondNature. Sitio web: <https://www.secondnature.com/>
- c) Amazon Web Services. Sitio web: <https://aws.amazon.com/>
- d) Docker. Sitio web: <https://www.docker.com/>

6. Algoritmos de (δ, γ) -Matching y sus Aplicaciones

Estudiante **Jesus Andres Rojas Montenegro***

Director *Juan Carlos Mendivelso Moreno***

Emails *jrojasmo@unal.edu.co, **jcmendivelsom@unal.edu.co

RESUMEN. La resolución eficiente de patrones de búsqueda es una rama importante en el campo de las Ciencias de la Computación. Específicamente, en contextos como secuencias musicales o biomoleculares, se busca identificar patrones permitiendo algunas discordancias pero que se siga manteniendo el parecido con el patrón. Para los textos y patrones formados por enteros, es posible medir la similitud entre dos cadenas usando dos tolerancias: δ y γ . La tolerancia δ establece el límite para la diferencia entre caracteres con el mismo índice, mientras que la tolerancia γ indica el máximo acumulado de diferencias permitidas en la suma de distancias.

El problema de encontrar patrones bajo esta medida de similitud entre cadenas se conoce como (δ, γ) -*matching*. En este trabajo de grado, se llevó a cabo una investigación e implementación de la mayoría de los algoritmos existentes para abordar este problema. Además, se presentó un nuevo algoritmo que usa Interval Trees, el cual fue evaluado y comparado en términos de rendimiento junto con los demás algoritmos, a través de diversos experimentos que varían distintos parámetros.

En particular, resulta que uno de los mejores algoritmos, (δ, γ) -Tuned Boyer-Moore, parte de una idea simple en comparación a otras ideas que usan estructuras mucho más interesantes y obtuvo los mejores resultados experimentales. Por otra parte, el algoritmo Backward Scan resulta bueno para patrones y δ grandes. Para δ pequeños, los otros también tuvieron un buen desempeño. Además, se exploraron algunas aplicaciones del problema de (δ, γ) -*matching*, en donde se refleja el hecho que este se adapta a contextos varios. Como producto final quedan todos los algoritmos estudiados implementados en C++ para su uso.

7. Procesos ETL y análisis de datos

Estudiante **Juan David Sanchez Murcia***

Director *Humberto Sarria Zapata***

Emails *jusanchezmu@unal.edu.co, **hsarriaz@unal.edu.co