



DEPARTAMENTO DE
**INGENIERÍA
INFORMÁTICA**

[A.R.R.A.Y]

APPLIED RESEARCH DAY

Tercera Jornada de Investigación Aplicada - DIINF 2024

23 de octubre 2024



CONTENIDOS

APPLIED RESEARCH DAY

Introducción	4
Dos miradas	4
Investigación	5
Vinculación	6
Hackatón	6
Reconocimiento	6



INF. APLICADA EN BIOLOGÍA Y MEDICINA

Acosta-Méndez, Christian. Desarrollo de aplicación web para el análisis filogenético de datos paleontológicos.	7
Aliaga-Rojas, Sebastián – Enfoque integrado para la representación del conocimiento y el razonamiento en bioinformática.	8

Arias-Panez, Israel – Relación entre los espacios de decisión y objetivo en el problema de inferencia de redes filogenéticas MO. 9

Castro-Castro, Sofía – Diseño de un índice de opt. basado en contexto biológico para un algoritmo de selección de características MO. 10

González-Hurtado, Benjamín – Paquete para caracterizar los resultados de algoritmos MO en el agrupamiento de genes. 11

Huincahue-Azócar, Ítalo – Paquete de extensión de MO para la captura, visualización y análisis de datos de rastreo ocular y fisiológicos. 12

Landero-Sepúlveda, Pamela – Conjuntos alternativos como estrategia de BL para selección de características en expresión génica. 13

Osswald-Dinamarca, Dominique – Evaluación de la ET en hospitales chilenos: análisis comparativo antes y durante periodo de pandemia. 14

Agradecemos profundamente a cada académico, académica, funcionario, funcionaria, administrativo y estudiante que hizo posible esta jornada. Su dedicación y entusiasmo hicieron de este encuentro una experiencia significativa para toda nuestra comunidad.



CONTENIDOS

Pérez-Sanhueza, Carlos. – 15
Calificación automática de biopsias HER2 de cáncer de mama utilizando aprendizaje autosupervisado.

Rivera-Alarcón, Claudia – 16
Uso de modelo LLM como caracterizador de coherencia biológica en clustering MO de datos de expresión génica.

Roa-Rodríguez, Joakin – 17
Simulación de la actividad neuronal de los ganglios basales.

Rousseau-Riveros, Francisco – 18
GS de biopsias de cáncer de mama HER2 usando aprendizaje supervisado basado en anotaciones.



Salas-Valenzuela, Bryan. – 19
Dominio para clasificación de espermatozoides humanos.

Valenzuela-Silva, Maximiliano. – 20
Evaluación de la ET en sistemas de salud entre 2015 y 2020: análisis comparativo y efecto de la pandemia.

Vásquez-Salgado, Javiera. – 21
Análisis evolutivo y de genes asociados TORC1 en *S. cerevisiae* mediante IA



 INF. APLICADA EN EDUCACIÓN

Chourio-Acevedo, Luz – 22
Modelo educativo con asistente virtual para búsqueda informacional.

Muñoz-Montecinos, Claudio – 23
Asistente virtual para fortalecer competencias de búsqueda en línea.

Toro-Flores, Luis – 24
Simulando el rol de gestor de riesgos en ciberseguridad: Un enfoque educativo mediante un videojuego RPG.

 INF. APLICADA EN LA SOCIEDAD

Céspedes Arnella, Florencia – 25
Reconocimiento de pánico en la voz con técnicas de computación afectiva y reconocimiento de voz.

Corral-Rodríguez, Luis. – 26
Holografía acústica en la superficie de fuentes sonoras mediante dispositivos de bajo costo.

Agradecemos profundamente a cada académico, académica, funcionario, funcionaria, administrativo y estudiante que hizo posible esta jornada. Su dedicación y entusiasmo hicieron de este encuentro una experiencia significativa para toda nuestra comunidad.



CONTENIDOS

Silva-Sanguinetti, Stephan – 27
Preprocesamiento con GPU usando polinomios ortogonales y multiprocesado para imágenes.

Vilches-Urrutia, Ángel – Sistema de software para la creación de datasets en ambientes smarthome. 28



 INF. APLICADA EN LA INDUSTRIA

Arancibia-Vásquez, Ricardo – 29
Predicción de la calidad del café mediante GLM.

Ballesteros-Ramírez, Héctor – App para automatizar procesamiento tributario desde el SII. 30

Donoso-Ocares, Guillermo – 31
Asignación de recursos en redes móviles con aprendizaje por refuerzo.

García-Guajardo, Camila – 32
Predicción de fallas en ventiladores mineros con redes neuronales.

Rickmers-Blamey, Ian – GripRoute: innovación tecnológica para escaladores y gimnasios. 33

 INF. APLICADA EN LA CIENCIA

Argomedo-Madrid, Ariel – 34
Parámetros de discos proto planetarios utilizando aprendizaje automático.

Calderón-Rivas, Daniel – 35
Optimización del entrelazamiento cuántico en sistemas grandes mediante cómputo de alto rendimiento.

López-Huarnez, Benjamín – 36
Explicabilidad orientada a objetivos en métodos de aprendizaje por refuerzo basados en gradiente de estrategias.

Loyola-Jara, Bastián – Framework evolutivo para el diseño de algoritmos neuromórficos. 37

Pineda-Moya, José – Solución neuromórfica al problema de bandidos contextual mediante un modelo de los ganglios basales. 38

Sepúlveda-Ballesteros, Nicolás – Uso de técnicas de machine learning para la detección de anomalías a partir de imágenes astronómicas. 39

Valero-Croma, David – Uso de polinomios ortogonales discretos para el preprocesamiento en la síntesis de imágenes. 40

Vergara-Sepúlveda, Leo – 41
Reconstrucción de imágenes con datos polarizados mediante algoritmos de Regularized Maximum Likelihood.

Vidal-Bormuth, Alma – Modelo de deep learning para estimar la dinámica en discos protoplanetarios. 42

Agradecemos profundamente a cada académico, académica, funcionario, funcionaria, administrativo y estudiante que hizo posible esta jornada. Su dedicación y entusiasmo hicieron de este encuentro una experiencia significativa para toda nuestra comunidad.

ROBERTO
GONZÁLEZ IBÁÑEZ



DIRECTOR
DEPARTAMENTAL

APPLIED RESEARCH DAY ARRAY

Desde el Departamento de Ingeniería Informática (DIINF) de la Universidad de Santiago de Chile, entendemos la investigación como parte inseparable del proceso formativo. En ese espíritu, **la tercera Jornada de Investigación Aplicada 2024,**

es una instancia clave donde nuestras y nuestros estudiantes comparten sus avances con la comunidad, fortaleciendo así una cultura académica orientada a la aplicación del conocimiento con impacto.

DOS MIRADAS...

Durante la jornada, diferentes relatores y relatoras compartieron perspectivas que cruzaron la crítica conceptual con la experiencia técnica. **Gabriela Arriagada Bruneau**, desde la filosofía, problematizó la inteligencia artificial a partir de sus tensiones éticas, mientras **Juande Santander-Vela**, con trayectoria en astrofísica, abordó los desafíos estratégicos y computacionales de la modernización del radiotelescopio ALMA.



Desde la práctica y el entorno empresarial, **Juan Manuel Barrios** **relató** cómo Impresee evolucionó desde una investigación académica hacia una solución tecnológica para el comercio electrónico. Por su parte, **Jorge González Paz** abordó los desafíos actuales de la Arquitectura Digital, destacando el impacto de la inteligencia artificial generativa en grandes organizaciones.





ESTUDIANTADO

“El protagonismo del estudiantado en la investigación es clave para una universidad conectada con su entorno...”

INVESTIGACIÓN

Más de **36** estudiantes de pregrado, magíster y doctorado presentaron investigaciones aplicadas junto a académicos y académicas del DIINF, a través de pósters, charlas y demos, abarcando cinco áreas clave de aplicación.



En **biología y medicina** (15), se propusieron soluciones desde la bioinformática e ingeniería biomédica. En **educación** (3), los trabajos desarrollaron herramientas para fortalecer habilidades informacionales, integrar sensores en videojuegos y simular escenarios de ciberseguridad.

En el ámbito de la **sociedad** (4), se trabajó en reconocimiento emocional, visualización acústica y análisis de datos públicos. En **industria** (5), los proyectos abarcaron minería, comercio electrónico, ventas, producción de café y automatización tributaria.

Por último, en **ciencia** (9), se abordaron temas como astroinformática, computación cuántica, aprendizaje por refuerzo y representación de conocimiento biológico.

El detalle de los trabajos se encuentra en la última sección.



VINCULACIÓN

La jornada contó con la participación de diversos **equipos institucionales** que fortalecieron el vínculo con la comunidad: la Asociación de Informáticos UTE-USACH (A.G.), la Biblioteca USACH, acompañamiento, equipo de CITIAPS y un espacio de apoyo de la Vicerrectoría de Investigación, Innovación y Creación (VRIIC).

HACKATÓN

Diferentes grupos de estudiantes participaron en una intensa jornada de **hackatón** organizada por **NTT DATA** y **DIINF-USACH**, donde desarrollaron propuestas innovadoras para transformar la búsqueda de talento. La actividad concluyó con la premiación al equipo ganador. Paralelamente, se efectuó un taller abierto de Ciberseguridad.



RECONOCIMIENTO

Al finalizar el evento se **reconocieron** los tres **mejores** trabajos presentados en la jornada, destacando a **Alma Vidal-Bormuth** por su modelo de deep learning en astronomía, **Luz Chourio-Acevedo** por su propuesta educativa con asistente virtual, y **Ángel Vilchez-Urrutia** por su sistema aplicado a entornos smarthome.

DESARROLLO DE APLICACIÓN WEB PARA EL ANÁLISIS FILOGENÉTICO DE DATOS PALEONTOLÓGICOS

CHRISTIAN ACOSTA-MÉNDEZ & MANUEL VILLALOBOS-CID

Este proyecto desarrolla una plataforma web basada en el pipeline de Concha (2023) para generar árboles filogenéticos de consenso sin politomías. Utilizando R y Shiny, busca facilitar su uso en paleontología, mejorando la precisión evolutiva y superando limitaciones de métodos tradicionales. Será evaluada por su usabilidad, robustez y eficacia.

Este proyecto se enfoca en el desarrollo de una plataforma web para facilitar la aplicación del pipeline propuesto por Concha (2023), orientado a la obtención de árboles filogenéticos de consenso sin las complicaciones que presentan las politomías, especialmente en el análisis de secuencias morfológicas. La motivación surge de la necesidad de superar las limitaciones de las técnicas de consenso tradicionales, que suelen generar resultados ambiguos, dificultando la interpretación de relaciones evolutivas entre especies prehistóricas.

El problema principal radica en la falta de herramientas accesibles para que paleontólogos e investigadores puedan aplicar este pipeline, que permite seleccionar el árbol medoide más representativo a partir de un conjunto de árboles generados bajo el criterio de máxima parsimonia. Aunque esta metodología ha sido validada conceptualmente, aún no cuenta con una implementación práctica y amigable que facilite su uso en entornos reales de investigación.

La solución propuesta consiste en desarrollar una plataforma web intuitiva y robusta, basada en una metodología de desarrollo iterativa e incremental que permita incorporar retroalimentación progresiva. La implementación se realiza utilizando herramientas como RStudio, GitHub y el entorno Shiny, aprovechando la versatilidad de R para el manejo de datos filogenéticos.

Se espera obtener una plataforma funcional que permita a los investigadores generar, analizar y visualizar árboles filogenéticos de manera más precisa, minimizando la ambigüedad asociada a las politomías. La plataforma será evaluada a través de pruebas de usabilidad y robustez, y sus resultados serán comparados con estudios filogenéticos previos para validar su eficacia y utilidad en el ámbito paleobiológico.



ENFOQUE INTEGRADO PARA LA REPRESENTACIÓN DEL CONOCIMIENTO Y EL RAZONAMIENTO EN BIOINFORMÁTICA

SEBASTIÁN ALIAGA-ROJAS & MARIO INOSTROZA-PONTA

Este trabajo propone unificar datos biológicos ómicos dispersos mediante una representación del conocimiento basada en Inteligencia Artificial (KRR). Al consolidar diversas bases de datos y automatizar los análisis con herramientas avanzadas de razonamiento, busca descubrir relaciones novedosas y mejorar la eficiencia en la comprensión de procesos biológicos complejos.

El conocimiento biológico es diverso, dinámico y está marcado por una alta incertidumbre y escalabilidad, disperso en numerosas bases de datos especializadas. Esta diversidad surge de la clasificación realizada por expertos, quienes organizan el conocimiento según su naturaleza y las posibilidades de análisis, dando lugar a categorías como datos metabólicos, genómicos y proteómicos. Estas categorías, conocidas como datos ómicos, reflejan las diferentes tecnologías disponibles para explorar procesos moleculares específicos y condiciones generales de los organismos. Sin embargo, esta segmentación del conocimiento representa un desafío significativo para los investigadores que buscan realizar análisis integradores, ya que dificulta la comprensión holística de los procesos biológicos al combinar múltiples perspectivas ómicas en un solo estudio.

Esta tesis doctoral propone una solución desde el campo de Representación y Razonamiento del Conocimiento (Knowledge Representation and Reasoning, KRR), una rama de la Inteligencia Artificial. La propuesta consiste en desarrollar una nueva representación unificada del conocimiento que consolide diferentes tipos de información biológica provenientes de diversas bases de datos. Al integrar este conocimiento en un marco único, la tesis busca automatizar los análisis mediante herramientas avanzadas de razonamiento inherentes a KRR. Estas herramientas, que emplean métodos lógicos o vectoriales, permitirán descubrir relaciones nuevas o poco evidentes dentro de los datos. Este enfoque promete mejorar el análisis de datos biológicos, haciéndolo más eficiente al pasar de procesos manuales a procesos automatizados.



RELACIÓN ENTRE LOS ESPACIOS DE DECISIÓN Y OBJETIVO EN EL PROBLEMA DE INFERENCIA DE REDES FILOGENÉTICAS MULTI OBJETIVO

ISRAEL ARIAS-PANEZ & MANUEL VILLALOBOS-CID

Este estudio analiza la relación entre el espacio de soluciones y el espacio objetivo en la inferencia de redes filogenéticas. Utilizando clustering y métricas topológicas, se demuestra que no existe una correlación directa entre ambos espacios, lo que destaca la necesidad de nuevas estrategias de optimización que consideren la diversidad estructural.

Las redes filogenéticas permiten representar relaciones evolutivas complejas que los árboles filogenéticos tradicionales no logran capturar. Estas redes se infieren mediante algoritmos de optimización, y en particular se ha explorado el uso de técnicas de optimización multiobjetivo. No obstante, estas estrategias enfrentan un desafío fundamental: el desconocimiento de la relación entre el espacio de soluciones —es decir, las topologías de redes generadas— y el espacio objetivo, conformado por los criterios de optimización utilizados. Esta desconexión puede estar provocando la exclusión de soluciones relevantes durante el proceso de búsqueda.

Las estrategias actuales se enfocan principalmente en optimizar criterios definidos en el espacio objetivo, sin garantizar una adecuada diversidad en el espacio de soluciones. Esto limita la exploración y podría inducir sesgos en la inferencia filogenética, ya que redes con estructuras muy distintas pueden alcanzar puntajes similares, y viceversa. Comprender la relación entre ambos espacios es clave para mejorar la calidad y representatividad de las soluciones generadas.

Este trabajo investiga dicha relación utilizando técnicas de clustering para analizar múltiples redes filogenéticas generadas mediante operadores de reordenamiento y mutación. Las redes fueron evaluadas según métricas topológicas y criterios objetivos como la verosimilitud y la parsimonia. Luego, se aplicaron algoritmos de agrupamiento para identificar patrones compartidos y se utilizó el índice de Jaccard para medir la coincidencia entre agrupamientos en ambos espacios.

Los resultados revelan que, en la mayoría de los casos, no existe una correspondencia directa entre el espacio topológico y el espacio objetivo: redes morfológicamente diferentes pueden mostrar rendimientos similares, mientras que redes similares pueden obtener puntuaciones distintas. Estos hallazgos evidencian que las estrategias actuales de optimización podrían estar sesgadas, y subrayan la necesidad de diseñar nuevos enfoques que integren explícitamente la diversidad topológica como un objetivo más en el proceso de optimización.

Se concluye que el espacio de decisiones no está alineado con el espacio objetivo, lo que sugiere que muchas soluciones con potencial evolutivo relevante podrían estar siendo descartadas. Esto abre la puerta al desarrollo de nuevas estrategias de optimización multiobjetivo que logren una exploración más balanceada, integrando simultáneamente criterios de calidad y diversidad estructural.



DISEÑO DE UN ÍNDICE DE OPTIMIZACIÓN BASADO EN CONTEXTO BIOLÓGICO PARA UN ALGORITMO DE SELECCIÓN DE CARACTERÍSTICAS MULTIOBJETIVO

SOFÍA CASTRO-CASTRO & MARIO INOSTROZA-PONTA

Se propone un enfoque multiobjetivo para la selección de características en datos de expresión génica, incorporando conocimiento biológico de bases como GO, KEGG y DisGeNet. Se diseña un índice contextual que guía el proceso de selección, mejorando la coherencia funcional y el rendimiento estadístico frente a métodos tradicionales sin información biológica.

La bioinformática es una disciplina que combina la biología y la informática para aplicar herramientas computacionales y métodos analíticos que permiten adquirir, comprender e interpretar información funcional de naturaleza biológica. Esta información varía según la ciencia ómica abordada; por ejemplo, la transcriptómica se enfoca en los patrones de expresión génica a partir del ARN mensajero, generando datos de alta dimensionalidad. Este tipo de datos plantea el desafío de seleccionar características relevantes, lo cual permite reducir la dimensionalidad de entrada y mejorar el rendimiento de los algoritmos de análisis PÓSTERiores.

En los últimos años, el análisis de datos de expresión génica ha evolucionado hacia enfoques más integradores, incorporando información proveniente de diversas bases de conocimiento biológico como Gene Ontology (GO), KEGG y DisGeNet. Si bien este conocimiento ha sido ampliamente utilizado en tareas de clasificación, asociación y aprendizaje automático, su incorporación en procesos de selección de características ha sido escasa.

En consecuencia, los métodos tradicionales que dependen únicamente de datos experimentales tienden a generar soluciones que no necesariamente reflejan la relevancia biológica de los genes seleccionados.

Este proyecto propone el diseño de un índice biológico que integre conocimiento contextual de múltiples bases de datos, con el fin de incorporarlo en los algoritmos de selección de características de forma contextualizada. Esta estrategia permitirá seleccionar subconjuntos de genes no solo desde una perspectiva estadística, sino también biológicamente coherente y funcionalmente relevante.

La metodología adoptada transformará el problema en una tarea de optimización multiobjetivo (MOO), permitiendo maximizar simultáneamente el rendimiento computacional y la coherencia biológica. La validación de los resultados incluirá: revisión sistemática de la literatura para verificar la consistencia biológica, evaluación estadística del rendimiento, validaciones funcionales con bases de datos especializadas, y comparación con métodos convencionales sin contexto biológico.



PAQUETE PARA CARACTERIZAR LOS RESULTADOS DE ALGORITMOS MULTIOBJETIVO EN EL AGRUPAMIENTO DE GENES

BENJAMÍN GONZÁLEZ-HURTADO & MARIO INOSTROZA-PONTA

Este proyecto desarrolla una librería en Python para comparar soluciones de clustering multiobjetivo en datos de expresión génica. Analiza la Frontera de Pareto usando medidas cuantitativas (índice de Jaccard) y cualitativas (enriquecimiento biológico con índice de Wang), facilitando la interpretación de agrupamientos y su relevancia biológica. Se evalúa con datos de CuMiDa.

El clustering multiobjetivo aplicado a datos de expresión génica es una técnica avanzada en bioinformática que permite agrupar genes según patrones de comportamiento observados en diversas condiciones experimentales, optimizando simultáneamente múltiples criterios de calidad. A diferencia de métodos clásicos como el hierarchical clustering, este enfoque utiliza algoritmos metaheurísticos para explorar eficientemente el espacio de soluciones, generando un conjunto de agrupamientos igualmente válidos que conforman la Frontera de Pareto (FP), donde ninguna solución domina a otra en todos los objetivos.

Si bien este conjunto de soluciones ofrece un panorama más completo y flexible, también plantea desafíos importantes a la hora de interpretarlo, debido al volumen y diversidad de resultados generados. Para abordar esta problemática, el presente proyecto propone el desarrollo de una librería en Python orientada a comparar y caracterizar los resultados de agrupamiento generados por algoritmos multiobjetivo en datos de expresión génica. Esta herramienta permitirá realizar evaluaciones exhaustivas, tanto cuantitativas como cualitativas, sobre las soluciones contenidas en la FP.

El enfoque cuantitativo incluye el uso de índices de similitud como Jaccard, que permiten medir el grado de coincidencia entre agrupamientos, mientras que el análisis cualitativo se basa en el enriquecimiento biológico de los grupos mediante información proveniente de Gene Ontology, evaluada con el índice de Wang. La herramienta será evaluada usando datasets del repositorio CuMiDa y ofrecerá visualizaciones y estructuras de datos que faciliten la comprensión de las similitudes y diferencias entre soluciones.

Como resultado, se espera que esta librería proporcione un recurso de libre acceso que permita a investigadores analizar, interpretar y seleccionar agrupamientos de alta calidad biológica, facilitando futuros estudios en expresión génica mediante técnicas de clustering multiobjetivo.



PAQUETE DE EXTENSIÓN DE MULTIMODAL OBSERVER PARA LA CAPTURA, VISUALIZACIÓN Y ANÁLISIS DE DATOS DE RASTREO OCULAR Y FISIOLÓGICOS

ÍTALO HUINCAHUE-AZÓCAR & ROBERTO GONZÁLEZ-IBÁÑEZ

Multimodal Observer es un software extensible para capturar, visualizar y analizar datos multimodales. Este trabajo desarrolló extensiones para integrar dispositivos de eye tracking (Gazepoint3HD) y manejo de datos fisiológicos (BiosignalPlux), utilizando la metodología RAD. Las evaluaciones confirmaron la efectividad de las extensiones para ampliar las capacidades de MO.

Multimodal Observer (MO) es un software desarrollado por la Universidad de Santiago de Chile, diseñado para la captura, visualización y análisis de datos provenientes de diversas fuentes, como la actividad del mouse, teclado, sensores de rastreo ocular, dispositivos de manejo de datos fisiológicos, entre otros (datos multimodales). Estos tipos de datos se utilizan frecuentemente en procesos de investigación como análisis conductual, interacción humano-computador, estudios de actividad fisiológica, marketing, entre otros.

Aunque MO se presenta como una herramienta que apoya los procesos de investigación mediante la captura, visualización y análisis de datos, la aparición de nuevos dispositivos y sensores para satisfacer necesidades específicas de estudio ha dejado algunos de ellos sin soporte en la plataforma. Dado que MO tiene la capacidad de ser extensible, este trabajo propone el desarrollo de extensiones para incorporar herramientas que permitan la captura, visualización y análisis de datos provenientes de dos dispositivos: el rastreador ocular Gazepoint3HD y el dispositivo de manejo de datos fisiológicos BiosignalPlux.

El desarrollo de este proyecto se guió por la metodología Rapid Application Development (RAD), utilizando un proceso iterativo de prototipado y refinamiento hasta obtener el producto final. Se creó un conjunto de extensiones enfocadas en la captura, visualización y análisis, las cuales fueron evaluadas dentro de MO. Los resultados confirmaron que la capacidad del sistema de plugins para extenderse es efectiva y adecuada para integrar nuevos dispositivos.



EVALUACIÓN DE CONJUNTOS DE CARACTERÍSTICAS ALTERNATIVOS COMO ESTRATEGIA DE BÚSQUDA LOCAL PARA EL PROBLEMA DE SELECCIÓN DE CARACTERÍSTICAS EN EXPRESIÓN GÉNICA

PAMELA LANDERO-SEPÚLVEDA & MARIO INOSTROZA-PONTA

Este trabajo explora el uso de conjuntos alternativos de características como estrategia de búsqueda local en metaheurísticas (SA, VNS, GA) para identificar biomarcadores genéticos en datos de expresión génica. Usando conjuntos de datos de cáncer del repositorio CUMIDA, se lograron soluciones de alta calidad y paneles de genes más pequeños y eficientes.

La identificación de biomarcadores genéticos es un proceso clave en el análisis de datos de expresión génica. Este proceso permite identificar posibles objetivos moleculares para el diseño de medicamentos o tratamientos para enfermedades como el cáncer y el Alzheimer, entre otras. Este problema presenta la particularidad de que el número de características a analizar está en el rango de miles, mientras que el número de muestras suele ser, como máximo, de uno o dos cientos. Además, las características pueden estar altamente correlacionadas entre sí, lo que genera miles de alternativas con calidad similar o aceptable para un conjunto seleccionado de características.

Este trabajo explora el uso de conjuntos alternativos de características como una estrategia de búsqueda local, implementada en tres metaheurísticas ampliamente conocidas: Recocido Simulado (Simulated Annealing, SA), Búsqueda Local de Vecindad Variable (Variable Neighborhood Local Search, VNS) y Algoritmo Genético (Genetic Algorithm, GA). Se utilizaron conjuntos de datos de expresión génica relacionados con el cáncer, obtenidos del repositorio público CUMIDA (<https://sbcb.inf.ufrgs.br/cumida>), para evaluar la propuesta.

Los resultados muestran que el uso de simLS (la estrategia de búsqueda local propuesta) permitió encontrar soluciones con mejor o similar calidad en comparación con no usarla dentro de una metaheurística. Además, en la mayoría de los casos, esta estrategia logró identificar paneles de genes más pequeños.



EVALUACIÓN DE LA EFICIENCIA TÉCNICA EN HOSPITALES CHILENOS: ANÁLISIS COMPARATIVO ANTES Y DURANTE PERIODO DE PANDEMIA DE COVID-19

DOMINIQUE OSSWALD-DINAMARCA & MANUEL VILLALOBOS-CID

Este estudio analiza la eficiencia técnica de hospitales públicos chilenos antes y durante la pandemia de COVID-19 usando DEA ajustado por Case Mix Index. Se incorporan todos los hospitales y una regresión Tobit para identificar factores asociados. Los resultados ofrecen evidencia clave para evaluar la gestión de recursos hospitalarios en contextos críticos.

La eficiencia técnica (ET) mide la capacidad de maximizar la producción utilizando la menor cantidad posible de insumos, sin generar desperdicio. En el contexto hospitalario, la ET es crucial para determinar si un hospital está empleando sus recursos de manera óptima. En la literatura, su evaluación se realiza comúnmente mediante el Análisis Envolvente de Datos (DEA), técnica que compara la eficiencia de distintas unidades evaluando la relación entre sus salidas (outputs) y entradas (inputs). Este método es especialmente adecuado en hospitales, donde la heterogeneidad de pacientes y la variabilidad en el volumen de atención complican las comparaciones directas.

Para mejorar la equidad en estas comparaciones, se incluye el Case Mix Index (CMI), que pondera la complejidad de los pacientes atendidos según los Grupos Relacionados por Diagnóstico (GRD). Estos grupos agrupan a los pacientes por diagnósticos y procedimientos similares, permitiendo ajustar las salidas del modelo en función de la severidad clínica, lo que facilita una comparación más precisa entre hospitales.

Este tipo de análisis es particularmente relevante para los gobiernos, sobre todo durante crisis sanitarias como la pandemia de COVID-19, que obligó a optimizar recursos hospitalarios de forma drástica. Durante ese periodo, se observaron niveles significativos de ineficiencia en la gestión hospitalaria en países de América y Europa, en contraste con mejores resultados en regiones de Asia y África. En el caso chileno, los estudios publicados entre 2004 y 2023 presentan importantes limitaciones, como la exclusión de algunos hospitales públicos, inconsistencias en el uso de valores GRD y la ausencia de datos del periodo pandémico.

Este proyecto busca cerrar esa brecha mediante un análisis DEA completo que incluya todos los hospitales públicos del país, con CMI actualizados, para comparar la eficiencia técnica antes y durante la pandemia de COVID-19. Además, se incorporará un análisis de regresión Tobit para identificar los factores determinantes de la eficiencia observada. El objetivo es evaluar si la administración hospitalaria fue adecuada durante la crisis sanitaria y proporcionar evidencia útil para enfrentar futuras emergencias en el sistema de salud.



CALIFICACIÓN AUTOMÁTICA DE BIOPSIAS HER₂ DE CÁNCER DE MAMA UTILIZANDO APRENDIZAJE AUTOSUPERVISADO

CARLOS PÉREZ-SANHUEZA & VIOLETA CHANG-CAMACHO

Este trabajo adapta modelos de aprendizaje autosupervisado (DINO + ViT) para la calificación automática de biopsias HER₂ de cáncer de mama con muestras chilenas. Evalúa aumento de datos, hiperparámetros y robustez. Los resultados muestran mejoras en precisión y consistencia frente a métodos clásicos, contribuyendo al avance de herramientas diagnósticas en histopatología.

En los últimos años, especialmente en el campo de la histopatología, se han logrado avances relevantes en el uso de métodos automatizados para analizar tejidos en tareas de clasificación y calificación. Sin embargo, la ambigüedad intra e inter patólogo al etiquetar imágenes sigue siendo un desafío. Esto motiva la evaluación de modelos para la calificación automática de biopsias HER₂ de cáncer de mama, utilizando muestras de pacientes chilenos y aplicando métodos de aprendizaje autosupervisado.

Uno de los principales obstáculos es la ausencia de un estado del arte específico para este tipo de imágenes y problema, además de la variabilidad de los protocolos de tinción entre laboratorios. Este trabajo plantea adaptar y evaluar un modelo entrenado en un dominio morfológicamente cercano para enfrentar esta tarea.

La solución se centra en evaluar el comportamiento y rendimiento del modelo desde múltiples perspectivas: impacto del aumento de datos, optimización de hiperparámetros, y comparación con métodos clásicos y de aprendizaje profundo. También se analiza la robustez del modelo como indicador de su versatilidad en escenarios reales.

La metodología sigue los fundamentos del método científico, formulando hipótesis a partir de una pregunta de investigación y validándolas experimentalmente. La propuesta utiliza el modelo DINO en combinación con la arquitectura Vision Transformer (ViT), lo que ofrece mejoras significativas respecto de métodos tradicionales. Se destaca una mayor consistencia y objetividad en la clasificación, factores esenciales en la práctica clínica de la histopatología.



USO DE MODELO LLM COMO CARACTERIZADOR DE COHERENCIA BIOLÓGICA EN CLUSTERING MULTIOBJETIVO DE DATOS DE EXPRESIÓN GÉNICA

CLAUDIA RIVERA-ALARCÓN & MARIO INOSTROZA-PONTA

Este trabajo desarrolla un enfoque innovador que integra modelo extenso de lenguaje con el algoritmo NSGA-II para agrupar genes según expresión génica y coherencia biológica. Los LLM, como Gemini IA y bioGPT, asignan índices de coherencia efectivos, mejorando los resultados en comparación con métodos tradicionales, con aplicaciones prometedoras en investigación biomédica.

Este estudio busca desarrollar un enfoque innovador para la agrupación de genes que integre la coherencia biológica mediante modelos extensos de lenguaje (LLM). Este método promete una comprensión más profunda de la genética y sus interacciones, ofreciendo nuevas oportunidades en la investigación biológica y biomédica.

El principal desafío consiste en usar un LLM como caracterizador biológico dentro de un algoritmo metaheurístico para agrupar genes con similitudes en expresión génica y coherencia biológica. Esto incluye validar los LLM contextualizados con conocimiento biológico e integrarlos en el algoritmo NSGA-II.

El enfoque propuesto se desarrolla en dos fases: primero, evaluar la viabilidad de los LLM como caracterizadores biológicos; y segundo, integrarlos en el algoritmo NSGA-II, utilizando como funciones objetivo una matriz de expresión génica y el conocimiento biológico proporcionado por los LLM. Este método busca superar los resultados de los enfoques tradicionales.

Se prueban LLM como Gemini IA de Google y bioGPT de Microsoft, con resultados preliminares prometedores, validados con referencias bibliográficas. El índice de coherencia biológica, calculado por los LLM, varía según el conjunto y nivel de coherencia evaluado.

El estudio adopta un enfoque cuantitativo correlacional para analizar la relación entre variables independientes (datos de expresión génica y metaheurística) y dependientes (calidad del agrupamiento de genes, medida por coherencia biológica y valores de hipervolumen).

El desarrollo incluye construir un caracterizador biológico evaluando modelos LLM, diseñando conjuntos de entrenamiento, prompts y ajustando resultados. Este caracterizador se integra en NSGA-II para optimizar el rendimiento. Finalmente, los grupos se evalúan mediante enriquecimiento, pruebas estadísticas y visualización.

Los resultados preliminares confirman que los LLM pueden asignar índices de coherencia biológica que mejoran el agrupamiento en NSGA-II.



SIMULACIÓN DE LA ACTIVIDAD NEURONAL DE LOS GANGLIOS BASALES

JOAKIN ROA-RODRÍGUEZ & JAVIER BALADRÓN PEZOA

Este trabajo simula un experimento ideomotor mediante un modelo computacional de ganglios basales. Los resultados reproducen tiempos de respuesta observados en humanos, confirmando que asociaciones acción-efecto afectan la planificación motora. El modelo sugiere que el fenómeno ideomotor emerge de la organización jerárquica de los bucles córtico-basales, vinculando conducta y neuroanatomía.

Las teorías psicológicas formalizadas en modelos computacionales suelen abstraerse de la actividad neuronal, lo que dificulta establecer una conexión directa entre los fenómenos conductuales y los procesos neuronales subyacentes. Esta desconexión se debe en parte a la falta de consistencia con la neuroanatomía real. En los últimos años, la teoría ideomotora ha resurgido como un marco teórico prometedor, especialmente por su aplicabilidad en la comprensión de ciertas patologías, como el síndrome de Tourette. El núcleo de esta teoría plantea que las acciones se seleccionan en función de representaciones anticipatorias de sus consecuencias. Entre los diversos estudios empíricos que la respaldan, destacan los paradigmas de adquisición de efectos, los cuales han demostrado experimentalmente la bidireccionalidad entre la representación de una acción y la de sus efectos.

Dado su rol central en la planificación de acciones, los ganglios basales han sido propuestos como la base neuronal de la teoría ideomotora.

En este trabajo se presenta un modelo computacional que simula los múltiples bucles funcionales entre los ganglios basales y la corteza cerebral, y se demuestra cómo este modelo puede explicar los resultados obtenidos en experimentos de adquisición de efectos. Las simulaciones generadas muestran patrones de comportamiento comparables a los observados en estudios realizados con personas, permitiendo formular predicciones sobre la relación entre distintas estructuras del circuito y el fenómeno ideomotor.

El experimento simulado consiste en una fase inicial en la que los participantes aprenden a seleccionar una acción frente a un estímulo, tras la cual se presenta un estímulo irrelevante (como un tono agudo o grave). PÓSTERiormente, en la fase de prueba, se presenta dicho estímulo irrelevante antes de la selección de acción. Los resultados indican un aumento en el tiempo de respuesta cuando el estímulo es inconsistente con la asociación previamente aprendida entre acción y efecto.

Este mismo efecto fue replicado por el modelo neurocomputacional, donde también se observaron tiempos de respuesta más lentos en los ensayos inconsistentes. Estos resultados sugieren que el fenómeno ideomotor puede entenderse como una consecuencia natural de la organización jerárquica de los ganglios basales, aportando una conexión directa entre la teoría psicológica y su correlato neurobiológico.



GOLD-STANDARD DE BIOPSIAS DE CÁNCER DE MAMA HER₂ USANDO APRENDIZAJE SUPERVISADO BASADO EN ANOTACIONES DE MÚLTIPLES PATÓLOGOS

FRANCISCO ROUSSEAU-RIVEROS & VIOLETA CHANG-CAMACHO

Este estudio analiza el impacto de distintas funciones de agregación en modelos de múltiples instancias aplicados a la predicción del HER2 Score en biopsias IHC de cáncer de mama. Los resultados permiten identificar configuraciones que mejoran la precisión del diagnóstico, contribuyendo al desarrollo de herramientas más confiables para la evaluación clínica automatizada.

El cáncer de mama es una enfermedad que requiere detección temprana y diagnóstico preciso para permitir un tratamiento eficaz. En este contexto, los modelos de inteligencia artificial basados en aprendizaje automático, particularmente aquellos que emplean aprendizaje de múltiples instancias (MIL), han surgido como herramientas prometedoras para la evaluación de esta enfermedad. Estos modelos utilizan biopsias obtenidas mediante inmunohistoquímica (IHC) para predecir el HER2 Score, un indicador clave en el manejo clínico del cáncer de mama.

A pesar de los avances en el uso de modelos MIL, el impacto que tiene la elección de la función de agregación final —mecanismo por el cual se combinan las representaciones de las instancias individuales para tomar una decisión a nivel global— no ha sido ampliamente explorado. Esta función puede influir de forma significativa en la capacidad predictiva del modelo, especialmente en contextos médicos donde la precisión diagnóstica es crítica.

Este trabajo se enfoca en investigar cómo diferentes funciones de agregación afectan el desempeño de modelos MIL en la predicción del HER2 Score a partir de biopsias IHC de cáncer de mama. A través de un análisis comparativo, se busca identificar qué funciones permiten una representación más precisa de la información relevante contenida en las imágenes histológicas, contribuyendo a mejorar el rendimiento del modelo y la calidad del diagnóstico.

Los resultados de esta investigación pueden tener un impacto directo en la mejora de modelos computacionales aplicados al diagnóstico de cáncer de mama, ofreciendo nuevas perspectivas para el desarrollo de herramientas más confiables y efectivas en contextos clínicos.



ADAPTACIÓN DE DOMINIO PARA CLASIFICACIÓN MORFOLÓGICA DE ESPERMATOZOIDES HUMANOS

BRYAN SALAS-VALENZUELA & VIOLETA CHANG-CAMACHO

Este estudio mejora la clasificación automática de morfología espermática adaptando un modelo de deep learning entrenado en dominios similares. Utilizando el dataset SCIAN-MorphoSpermGS, técnicas de aumento, penalización de clases, transfer learning y optimización bayesiana, se alcanzó una exactitud del 72,7%, superando notablemente el 63% logrado por métodos anteriores.

La infertilidad afecta al 15% de las parejas a nivel mundial, con cerca del 30% de los casos relacionados con factores masculinos. En este contexto, el análisis morfológico de los espermatozoides es fundamental para evaluar la fertilidad masculina. Sin embargo, este análisis, cuando se realiza manualmente, es subjetivo, poco reproducible y propenso a errores diagnósticos que pueden derivar en tratamientos inadecuados. Los modelos automáticos existentes para esta tarea, aplicados al dataset SCIAN-MorphoSpermGS, han alcanzado una exactitud máxima del 63%, lo que evidencia una oportunidad de mejora en la clasificación automática.

Este trabajo propone la adaptación de un modelo de deep learning que ha mostrado un rendimiento superior (más del 98% de exactitud) en un dominio morfológico similar. Para ello, se entrenó el modelo con el dataset SCIAN-MorphoSpermGS utilizando validación cruzada de 10 pliegues.

Se aplicaron técnicas de aumento de datos enfocadas en clases minoritarias para abordar el desbalance de clases, y se modificó la función de pérdida para penalizar las clases mayoritarias. Además, se utilizó aprendizaje por transferencia con dos datasets de dominios morfológicos afines (glóbulos blancos y granos de arroz), y se ajustaron los hiperparámetros mediante optimización bayesiana con la librería Optuna.

Como resultado, el modelo adaptado alcanzó una exactitud del 72,7% en la clasificación de morfología espermática, superando significativamente el desempeño de los modelos reportados previamente con el mismo conjunto de datos.



EVALUACIÓN DE LA EFICIENCIA TÉCNICA EN SISTEMAS DE SALUD GLOBALES ENTRE 2015 Y 2020: ANÁLISIS COMPARATIVO Y EFECTO DE LA PANDEMIA

MAXIMILIANO VALENZUELA-SILVA & MANUEL VILLALOBOS-CID

Este estudio evalúa la eficiencia técnica de sistemas de salud en diversos países entre 2015 y 2020, utilizando DEA, índice de Malmquist y regresión Tobit. Se analiza el impacto de la pandemia y se identifican políticas asociadas a altos niveles de eficiencia, proporcionando evidencia clave para la mejora en la gestión sanitaria global.

La eficiencia técnica (ET) es un componente clave para evaluar cómo los sistemas de salud utilizan sus recursos con el fin de alcanzar los mejores resultados posibles. En este contexto, la ET se refiere a la capacidad de un sistema sanitario para maximizar la prestación de servicios de salud a partir de insumos específicos como personal, infraestructura y financiamiento. Gestionar los sistemas de salud con altos niveles de ET es fundamental para garantizar que los recursos limitados se empleen de forma efectiva, mejorando la calidad y accesibilidad de la atención. Comparar la ET entre países proporciona una valiosa oportunidad para identificar buenas prácticas, detectar áreas críticas de mejora y permitir que los responsables de políticas aprendan de experiencias ajenas.

Actualmente, existe una notoria escasez de estudios que evalúen la ET de los sistemas de salud desde una perspectiva global. Aunque hay investigaciones de carácter regional, no se han realizado análisis que abarquen un amplio rango de países con distintas condiciones geográficas, económicas y sociales, ni que consideren la evolución de la ET en años recientes, particularmente durante la pandemia de COVID-19.

Esta carencia limita la comprensión sobre cómo se han gestionado los recursos sanitarios en escenarios de crisis y cuáles estrategias han sido más eficaces.

Para abordar esta brecha, se propone un estudio comparativo internacional que evalúe la ET de los sistemas de salud entre 2015 y 2020. El objetivo es identificar qué países han gestionado de manera más eficiente sus recursos y analizar las políticas y características asociadas a un mejor desempeño. Asimismo, se estudiará el impacto de la pandemia sobre los niveles de eficiencia.

La metodología contempla el uso del Análisis Envolvente de Datos (DEA) para medir la ET, considerando diferentes retornos a escala y orientaciones del modelo. Además, se aplicará el índice de Malmquist para observar la evolución temporal, diferenciando entre cambios técnicos y tecnológicos. También se utilizará una regresión Tobit para evaluar la influencia de factores externos, complementado con análisis de sensibilidad y un estudio de sub-periodos para evaluar el efecto de la pandemia.



ANÁLISIS EVOLUTIVO Y FUNCIONAL DE GENES ASOCIADOS AL COMPLEJO TORC1 EN SACCHAROMYCES CEREVISIAE MEDIANTE TÉCNICAS DE INTELIGENCIA ARTIFICIAL

JAVIERA VÁSQUEZ-SALGADO & MANUEL VILLALOBOS-CID

Este estudio aplica aprendizaje automático para analizar la historia evolutiva de genes vinculados al complejo TORC1 en *S. cerevisiae*. A través de árboles filogenéticos y agrupamiento funcional, se busca comprender mejor su adaptación a condiciones específicas, optimizando su uso en procesos fermentativos con aplicaciones industriales y biotecnológicas.

Saccharomyces cerevisiae es un organismo modelo fundamental en biotecnología debido a su capacidad para fermentar azúcares, lo que lo convierte en una pieza clave en la producción de alimentos, bebidas alcohólicas y productos farmacéuticos. En países como Chile, donde estas industrias tienen alta relevancia, resulta esencial profundizar en el conocimiento de los mecanismos que optimizan su uso industrial.

Aunque se han identificado 131 genes vinculados al complejo TORC1 en cepas vínicas de *S. cerevisiae*, aún no se comprenden completamente los mecanismos por los cuales este complejo detecta y responde a los niveles de nitrógeno, lo que representa una barrera para la mejora de procesos fermentativos en que participa. Para abordar este problema, se propone el uso de inteligencia artificial, específicamente algoritmos de aprendizaje automático, con el fin de analizar la historia evolutiva de los genes asociados al complejo TORC1. Este enfoque permitirá identificar genes con trayectorias evolutivas similares y caracterizarlos funcionalmente, lo que contribuirá a una mejor comprensión de sus funciones biológicas e industriales.

La metodología contempla la organización de una base de datos con las 6016 secuencias de ORFs del genoma de 1011 cepas de *S. cerevisiae*, la construcción de árboles filogenéticos para cada uno de ellos mediante el método Neighbor Joining, y la comparación de sus topologías con un árbol general utilizando la distancia de Robinson-Foulds. PÓSTERiormente, se aplicarán técnicas de visualización bidimensional (PCS, t-SNE o MDS) para agrupar genes según sus similitudes evolutivas, seguidas de un análisis de enriquecimiento funcional con ShinyGO. Todo el proceso se llevará a cabo en colaboración con expertos para la interpretación de los hallazgos y su documentación.

Como resultado, se espera identificar genes del complejo TORC1 con patrones evolutivos relacionados, lo que permitirá caracterizarlos funcionalmente y entender cómo han evolucionado para responder a condiciones específicas. Este conocimiento podrá ser aplicado para optimizar el uso de *S. cerevisiae* en procesos industriales y biotecnológicos, aumentando su eficiencia y rendimiento.



MODELO EDUCATIVO TECNOLÓGICO Y PROCEDIMENTAL BASADO EN UN ASISTENTE VIRTUAL PARA DESARROLLAR COMPETENCIAS DE BÚSQUEDA DE INFORMACIÓN EN LÍNEA EN ESTUDIANTES CHILENOS

LUZ CHOURIO-ACEVEDO & ROBERTO GONZÁLEZ-IBAÑEZ

Este proyecto desarrolla un asistente virtual para mejorar las competencias de búsqueda de información en línea de escolares en ciencias naturales. Evaluado en escuelas de Santiago, el modelo demostró aumentar la eficacia y motivación de los estudiantes, fortaleciendo su autonomía y apoyando a los docentes en contextos educativos digitalizados.

En la actualidad, niños y adolescentes crecen en un entorno altamente tecnologizado, donde el uso de la Web como principal recurso para realizar tareas escolares ha aumentado drásticamente en los últimos años. Dado el intenso uso de Internet en entornos educativos, las competencias de búsqueda, evaluación y uso de la información (alfabetización informacional) se consideran parte fundamental de las destrezas para el aprendizaje.

Sin embargo, en una era donde abundan los contenidos en formato electrónico y el docente ha dejado de ser el único proveedor de información, los estudiantes enfrentan el desafío de desarrollar nuevas competencias para alcanzar el éxito tanto en la vida académica como en la personal. En este contexto, estudios previos han demostrado que el aprendizaje está directamente relacionado con los comportamientos que los estudiantes exhiben al buscar información en la Web. A pesar de este incremento en el uso de Internet, las competencias de investigación en línea de los estudiantes no siempre están bien desarrolladas.

Este trabajo presenta un modelo tecnológico y procedimental basado en un asistente virtual para fomentar el desarrollo de competencias de búsqueda de información en línea, contextualizado en el aprendizaje de ciencias en escolares.

Para ello, (1) se caracterizaron los comportamientos de búsqueda en línea de los estudiantes, (2) se diseñó un asistente virtual para asistir los procesos de enseñanza-aprendizaje a partir de estos comportamientos en tareas de ciencias naturales, y (3) se llevó a cabo un experimento en escuelas particulares y municipales de Santiago para evaluar la efectividad del modelo propuesto en el desarrollo de competencias informacionales en este dominio.

Los resultados incluyen (1) un modelo tecnológico y procedimental y (2) una plataforma informática orientada a mejorar las competencias de búsqueda en línea de los estudiantes. Los hallazgos demuestran que las estrategias implementadas no solo mejoran significativamente estas competencias, permitiendo a los estudiantes encontrar información de manera más efectiva para sus tareas, sino que también generan una mayor motivación en comparación con las estrategias tradicionales. El modelo propuesto tiene el potencial de mejorar los resultados de aprendizaje y de apoyar eficazmente a los docentes en su labor formativa, al tiempo que fomenta el desarrollo autónomo de habilidades informacionales en los estudiantes.



DESARROLLO DE SENSORES PARA BGAMES Y SU APLICACIÓN CON UN MOD PARA EL VIDEOJUEGO TERRARIA

CLAUDIO MUÑOZ-MONTECINOS & ROBERTO GONZÁLEZ-IBÁÑEZ

Este proyecto amplía el framework bGames mediante sensores personalizados y un mod para Terraria, integrando actividades reales a la experiencia de juego. Usando metodología RAD, se desarrollaron herramientas para monitorear aplicaciones, navegación y ejercicio. El mod recompensa acciones del mundo real con beneficios en el juego, promoviendo hábitos saludables.

Este proyecto se centra en la expansión del framework Blended Games (bGames), diseñado para equilibrar la experiencia de juego con actividades reales de tipo físico, social, cognitivo y afectivo. En respuesta a la creciente preocupación por el uso excesivo de videojuegos, se propone el desarrollo de un sistema que integre datos del mundo real a la experiencia lúdica a través de sensores personalizados y un mod para el popular videojuego Terraria.

El desafío principal radica en la limitada disponibilidad de sensores precisos para monitorear actividades fuera del entorno virtual, lo que restringe la adopción de bGames. Para superar esto, se propone una solución en dos partes: (1) la creación e integración de sensores para monitorear aplicaciones, cursos en línea, navegación y actividad física con Fitbit; y (2) el desarrollo de un mod para Terraria que recompense al jugador dentro del juego por sus actividades reales, incentivando así comportamientos saludables y balanceados.

El desarrollo se llevó a cabo siguiendo un enfoque inspirado en el modelo Rapid Application Development (RAD), utilizando herramientas como Python, Docker, GitHub, tModLoader y Visual Studio. Los resultados incluyen sensores funcionales y un mod operativo que conecta al jugador con el sistema bGames, permitiendo obtener beneficios dentro del juego según sus actividades externas.



SIMULANDO EL ROL DE GESTOR DE RIESGOS EN CIBERSEGURIDAD: UN ENFOQUE EDUCATIVO MEDIANTE UN VIDEOJUEGO RPG

LUIS TORO-FLORES & JUAN ITURBE-ARAYA

Este proyecto propone un videojuego RPG educativo basado en Scrum para fortalecer la enseñanza de ciberseguridad organizacional en estudiantes de ingeniería informática. Utiliza roles y competencias definidas por ENISA, abordando brechas formativas y promoviendo el aprendizaje activo y contextualizado conforme a los lineamientos del Modelo Educativo Institucional.

La gestión de la ciberseguridad ha adquirido creciente relevancia en las organizaciones. Según reportes de la empresa Entel, Chile enfrenta un volumen significativo de amenazas cibernéticas, con un costo anual estimado cercano a los 7.700 millones de dólares. En este contexto, han surgido iniciativas que promueven la enseñanza y comprensión de la ciberseguridad. Destaca la European Union Agency for Cybersecurity (ENISA), que busca establecer un entendimiento común, contribuir a políticas cibernéticas y aumentar la confiabilidad de productos y servicios mediante esquemas de certificación.

ENISA ha publicado un documento donde identifica tareas, competencias, habilidades y conocimientos clave para los profesionales del área. Este material orienta a los programas formativos en el desarrollo de capacidades, promoviendo además la armonización en educación y capacitación del talento en ciberseguridad.

A nivel nacional, la Universidad se guía por el Modelo Educativo Institucional (MEI), que orienta los procesos de mejora continua en la formación. Este modelo destaca la “Focalización en el Aprendizaje”, entendida como la conexión entre conocimientos previos y nuevos aprendizajes, mediante metodologías didácticas diversas, colaborativas y desafiantes. También subraya la necesidad de experiencias de aprendizaje que permitan aplicar lo aprendido en contextos reales.

En la malla de Ingeniería Informática existen asignaturas como “Ciberseguridad” y “Seguridad y Auditoría”, pero su cobertura sobre la gestión de la ciberseguridad en entornos organizacionales es limitada. Aunque la asignatura de ciberseguridad aborda contenidos relevantes, los estudiantes suelen desconocer el funcionamiento interno de las organizaciones, lo que genera una brecha formativa. Además, la alta matrícula impide realizar un seguimiento individual del aprendizaje.

Como solución, se propone el desarrollo de un videojuego tipo RPG basado en Scrum, en el que los estudiantes asuman uno de los roles definidos por ENISA. El juego se diseña en torno a las habilidades descritas en el documento, facilitando la adquisición de conocimientos de ciberseguridad organizacional de manera lúdica, contextualizada y alineada con el modelo educativo institucional.



PROPUESTA DE APOYO AL RECONOCIMIENTO DE PÁNICO EN LA VOZ POR MEDIO DE TÉCNICAS DE COMPUTACIÓN AFECTIVA Y RECONOCIMIENTO DE VOZ

FLORENCIA CÉSPEDES ARNELLA & EDMUNDO LEIVA-LOBOS

Se propone un sistema de inteligencia artificial para identificar emociones complejas en la voz, útil en contextos críticos como emergencias o procesos judiciales. Mediante representación dimensional, el modelo apoya a expertos en la evaluación emocional del hablante, mejorando precisión y comprensión en situaciones de alta tensión o estrés.

El reconocimiento de emociones en la voz ha estado tradicionalmente enfocado en identificar emociones básicas, pero las emociones complejas también forman parte esencial de la experiencia humana. Su detección se vuelve especialmente desafiante en contextos de alta presión, como llamadas de emergencia, interrogatorios judiciales o testimonios en situaciones traumáticas, donde el estado emocional del hablante debe ser evaluado con precisión por expertos.

Con el objetivo de asistir a estos profesionales y aumentar la fiabilidad del análisis emocional en escenarios críticos, se propone el desarrollo de un sistema de inteligencia artificial capaz de identificar y representar emociones complejas a través de la voz. Este sistema busca modelar las emociones en un espacio dimensional, capturando matices emocionales más sutiles que escapan a las categorías tradicionales. Así, no solo se facilitaría una mejor interpretación del estado afectivo del hablante, sino que también se brindaría una herramienta tecnológica de apoyo en situaciones donde la comprensión emocional precisa es crucial.

Este enfoque abre nuevas posibilidades para abordar de forma más efectiva el reconocimiento de estados emocionales intensos en entornos judiciales, médicos o de atención de emergencias, entregando soporte analítico complementario que puede ser determinante en la toma de decisiones y en la evaluación profesional de situaciones delicadas.



HOLOGRAFÍA ACÚSTICA EN LA SUPERFICIE DE FUENTES SONORAS MEDIANTE DISPOSITIVOS DE BAJO COSTO

LUIS CORRAL-RODRÍGUEZ & PABLO ROMÁN-ASENJO

Se desarrolla un dispositivo portátil y económico para medir potencia acústica y velocidad de partículas en fuentes de ruido. Utiliza micrófonos MEMS, procesamiento GPU y diseño 3D modular. Validado en cámara anecoica, entrega resultados comparables a sistemas comerciales, facilitando el estudio acústico bajo normativas ambientales y ocupacionales vigentes.

El estudio de la radiación sonora de fuentes de ruido es fundamental para la formulación de políticas de mitigación, especialmente considerando que Chile no es ajeno a este tipo de contaminación en ambientes residenciales y ocupacionales. La Organización Mundial de la Salud advierte sobre los efectos adversos del ruido, tales como el aumento del riesgo de enfermedades coronarias, hipertensión, tinnitus, déficit cognitivo, pérdida auditiva, trastornos del sueño y deterioro de la salud mental.

Las tecnologías más recientes para evaluar fuentes de ruido incluyen arreglos de micrófonos y cámaras acústicas, que permiten obtener holografías de la distribución de la velocidad de partículas en la superficie emisora. Esta información permite identificar las zonas de mayor incidencia en la radiación sonora. Además, estos sistemas permiten medir la potencia acústica, un parámetro clave para predecir los niveles de presión sonora según la normativa ambiental y ocupacional vigente.

Sin embargo, los dispositivos actuales suelen ser costosos y voluminosos, lo que limita su acceso y aplicabilidad. Por esta razón, se presenta una alternativa portátil y de bajo costo, compuesta por un soporte impreso en 3D, placas de circuito para micrófonos MEMS, un dispositivo para el procesamiento digital de señales, una cámara de profundidad estéreo y una tarjeta de desarrollo con unidad de procesamiento gráfico (GPU). El sistema es compatible con Windows, Linux y macOS, y sus archivos de diseño están disponibles públicamente, permitiendo modificaciones para distintas configuraciones de micrófonos.

Las pruebas realizadas en una cámara anecoica muestran resultados satisfactorios en la estimación de potencia acústica y en la obtención de holografías de la velocidad normal de partículas. El dispositivo permite definir la geometría del emisor y medir variables acústicas clave. Utilizando una librería optimizada para GPU de NVIDIA, se logra resolver el problema acústico inverso en tiempos significativamente reducidos.



PRE-PROCESAMIENTO DE ALTO RENDIMIENTO CON GPU Y MULTIPROCESO BASADO EN POLINOMIOS DISCRETOS ORTOGONALES PARA SÍNTESIS DE IMÁGENES

STEPHAN SILVA-SANGUINETTI & PABLO ROMÁN-ASENJO

El radiotelescopio ALMA utiliza 66 antenas para captar visibilidades y generar imágenes del universo temprano. Este trabajo propone optimizar el preprocesamiento mediante GPU y multiproceso, mejorando la precisión y eficiencia del método de polinomios discretos ortogonales frente al gridding. Simulaciones evaluarán su rendimiento con datos públicos de ALMA.

El radiotelescopio ALMA (Atacama Large Millimeter/submillimeter Array), ubicado en el Desierto de Atacama, es fundamental para investigar la formación de estrellas y el universo temprano. Con 66 antenas interferométricas, ALMA capta ondas de luz que representan muestras de la transformada de Fourier de la luminosidad estelar, conocidas como "visibilidades". Estas son recogidas en puntos distribuidos de manera irregular debido a la configuración de las antenas, por lo que requieren algoritmos de síntesis de imágenes, como CLEAN, para transformar los datos de Fourier y generar imágenes del objeto estelar observado. El avance en la precisión y eficiencia del procesamiento de estas imágenes es clave para el progreso de las investigaciones realizadas por ALMA.

Para procesar estas imágenes, es necesaria una cuadrícula regular, lo que introduce el método de gridding, que interpola los datos a una cuadrícula y facilita la aplicación de la transformada de Fourier inversa (IFT). A pesar de su utilidad, el gridding presenta desafíos como el aumento del ruido, el alto costo computacional y la pérdida de significancia estadística.

Recientemente, se ha propuesto una alternativa basada en polinomios discretos ortogonales, que mejora la precisión y reduce la varianza. Sin embargo, esta solución enfrenta problemas relacionados con la eficiencia computacional y la complejidad del código.

Este trabajo propone mejorar esta solución mediante un enfoque distribuido que aprovecha GPU y multiproceso para optimizar el preprocesamiento de datos, además de mejorar la legibilidad y el mantenimiento del código. Asimismo, se realizarán simulaciones con datos públicos de ALMA para evaluar los límites técnicos y el rendimiento de la nueva implementación en comparación con la solución actual y un software de gridding, considerando las características del servidor utilizado para las pruebas.



SISTEMA DE SOFTWARE PARA LA CREACIÓN DE DATASETS EN AMBIENTES SMARTHOME

ÁNGEL VILCHES-URRUTIA & JUAN ITURBE-ARAYA

El Internet de las Cosas expone hogares inteligentes a riesgos de seguridad. Este trabajo propone un software que genera datasets realistas para mejorar los Sistemas de Detección de Intrusos, permitiendo capturar y analizar datos en tiempo real. Desarrollado con Python y metodología Personal Scrum, cumple estándares ISO y mejora la seguridad digital.

El Internet de las Cosas (IoT) conecta dispositivos físicos a internet, facilitando el intercambio, análisis y almacenamiento de datos, lo que transforma los hogares en ambientes inteligentes (smart homes). No obstante, esta tecnología implica riesgos significativos para la privacidad y seguridad, ya que la recopilación de datos personales expone a los usuarios a posibles intrusiones. Aunque existen Sistemas de Detección de Intrusos, su efectividad está limitada por la falta de conjuntos de datos (datasets) que representen adecuadamente escenarios reales, dificultando la seguridad en hogares inteligentes.

El principal problema radica en que los hogares inteligentes, al estar cada vez más conectados mediante dispositivos IoT, se han convertido en objetivos atractivos para ciberdelincuentes. La efectividad de los Sistemas de Detección de Intrusos, considerados la primera y principal barrera de seguridad de estos dispositivos, está comprometida debido al uso de conjuntos de datos obsoletos y poco representativos para el entrenamiento y evaluación de estos modelos. Esto dificulta enfrentar adecuadamente las amenazas actuales y futuras, comprometiendo la seguridad digital.

Como solución, se propone un software para generar conjuntos de datos realistas que reflejen con precisión los comportamientos en entornos de red. Diseñado para expertos en ciberseguridad, permite capturar y analizar datos en tiempo real, adaptándose a diversas configuraciones y ofreciendo herramientas como respaldo de datos, manejo de grandes volúmenes de información y etiquetado de actividades para facilitar el entrenamiento y evaluación de sistemas de detección de intrusos.

El desarrollo se llevó a cabo utilizando la metodología Personal Scrum, adaptada para un único profesional, lo que facilitó la gestión del tiempo y la priorización de tareas mediante sprints iterativos. Python fue el lenguaje principal, complementado con Jira y Git para gestión y control de versiones.

La evaluación del software mostró altos niveles de satisfacción en usabilidad (84%) y fiabilidad (90%), cumpliendo con los estándares definidos por la Organización Internacional de Normalización (ISO). También se confirmó su robustez en distintos escenarios, permitiendo crear datasets adecuados para evaluar sistemas de detección de intrusos.



PREDICCIÓN DE LA CALIDAD DEL CAFÉ Y SUS ATRIBUTOS DETERMINANTES MEDIANTE MODELOS LINEALES GENERALIZADOS

RICARDO ARANCIBIA-VÁSQUEZ & MANUEL VILLALOBOS-CID

Este proyecto aplica modelos predictivos para evaluar la calidad del café de forma eficiente y accesible. Integrando cata, análisis químicos y machine learning, busca escalar la evaluación más allá de micro lotes, permitiendo a productores optimizar procesos, mejorar calidad y aumentar rentabilidad, incluso en zonas emergentes como Rapa Nui.

La industria del café constituye el sustento económico de millones de familias en todo el mundo. A pesar de generar más de 200 mil millones de dólares anualmente, los y las productoras reciben solo una fracción de este ingreso, enfrentando precios que en muchos casos no cubren los costos de producción. Esta desigualdad se ve agravada por las dificultades que enfrentan para mejorar la calidad de sus cosechas, ya que los métodos actuales de evaluación —como la cata profesional— son costosos y poco escalables, restringiéndose a pequeños lotes de producción.

El problema también afecta a nuevos productores que desean incorporarse a esta industria. Tal es el caso de la incipiente producción de café de especialidad en la Isla de Rapa Nui, donde, si bien se cuenta con condiciones climáticas propicias, falta experiencia y conocimiento técnico sobre el comportamiento agronómico del cultivo en este entorno.

Este proyecto propone el uso de modelos predictivos para identificar los factores que influyen en la calidad del café y predecir sus características de manera eficiente. Se combinan enfoques tradicionales como la cata con métodos avanzados de análisis químico y técnicas estadísticas y de machine learning, con el objetivo de facilitar una evaluación de calidad más rápida, económica y aplicable a volúmenes mayores de producción.

Con esta propuesta, se busca entregar a los productores herramientas que les permitan realizar ajustes informados en sus prácticas agrícolas y de procesamiento, mejorando la calidad de su producto final, accediendo a nuevos mercados, y aumentando así sus márgenes de ganancia. Además, se promueve una producción más sostenible y rentable, democratizando el acceso a sistemas de evaluación de calidad a nivel industrial.



DESARROLLO DE APLICACIÓN PARA AUTOMATIZAR PROCESAMIENTO DE INFORMACIÓN TRIBUTARIA PROVENIENTE DESDE EL SERVICIO DE IMPUESTOS INTERNOS

HÉCTOR BALLESTEROS-RAMÍREZ & MANUEL VILLALOBOS-CID

Este proyecto desarrolla una aplicación de código abierto para automatizar la extracción y análisis de información tributaria desde el SII. Apunta a mejorar la productividad de consultoras pequeñas y medianas, reduciendo errores manuales y tiempos de procesamiento, especialmente en periodos críticos. Se desarrolla bajo metodología RAD, validado con SEP Consultores.

Este proyecto tiene como objetivo desarrollar una aplicación para automatizar el procesamiento de información tributaria extraída desde el Servicio de Impuestos Internos (SII), respondiendo a los problemas que enfrentan pequeñas y medianas consultoras tributarias al realizar estos procesos de forma manual. Actualmente, estas tareas son lentas, propensas a errores y consumen una gran cantidad de tiempo, especialmente en periodos de alta demanda como la Operación Renta.

El proceso de extraer, organizar y analizar información desde la Carpeta Tributaria Electrónica (CTE) del SII resulta ineficiente cuando se realiza manualmente, lo que puede generar demoras, errores e incluso sanciones por parte del organismo fiscalizador. Las herramientas actuales –como APIs de terceros o los propios servicios del SII– presentan limitaciones técnicas y económicas, dificultando el acceso a soluciones robustas para muchas empresas.

Frente a este escenario, el proyecto propone una aplicación que automatice la extracción y procesamiento de información tributaria desde el SII. Esta generará reportes con indicadores clave y asientos contables que facilitarán la toma de decisiones, permitiendo una gestión tributaria más eficiente. El desarrollo se llevará a cabo siguiendo la metodología RAD (Rapid Application Development), mediante prototipos iterativos validados en conjunto con SEP Consultores.

Además, la solución será de código abierto, lo que permitirá que otras empresas puedan personalizarla y adaptarla a sus propias necesidades, contribuyendo a un ecosistema colaborativo de herramientas para la gestión tributaria. Se espera que la aplicación reduzca de manera significativa los tiempos de procesamiento y los errores humanos, mejorando la productividad general de las empresas consultoras.



ASIGNACIÓN DE RECURSOS BAJO INCERTIDUMBRE EN EL PARADIGMA DE REDES MÓVILES: UN ENFOQUE DE APRENDIZAJE POR REFUERZO

GUILLERMO DONOSO-OCARES, KAREL TOLEDO-DE LA GARZA & JAVIER BALADÓN-PEZOA

Este trabajo propone usar drones como estaciones base móviles optimizadas con aprendizaje por refuerzo para maximizar la calidad de señal en redes dinámicas. Entrenados con Deep Q-Network y funciones de recompensa, los drones superaron a antenas estáticas. El mejor desempeño lo logró un agente que consideró el diferencial de throughput promedio entre estados.

En el mundo digital actual, la conexión a redes de telecomunicaciones es esencial, especialmente en áreas densamente pobladas o eventos masivos, donde la capacidad de las redes puede saturarse. Las antenas estáticas no se adaptan dinámicamente a los cambios en la distribución de usuarios, resultando en una calidad de servicio deficiente. Esto plantea la necesidad de estrategias en tiempo real que optimicen la cobertura y la capacidad del canal de comunicaciones.

Se propone utilizar un dron como estación base móvil, capaz de ajustar su posición para maximizar la cobertura y mejorar la calidad de señal para usuarios en movimiento. Diseñar estas estrategias de posicionamiento es un desafío, ya que el dron debe adaptarse a entornos dinámicos para maximizar la calidad de señal de un grupo variable de usuarios.

El proyecto utiliza algoritmos de aprendizaje por refuerzo para que el dron aprenda estrategias óptimas de cobertura y calidad de señal. Entrenado en un entorno simulado, donde usuarios y drones se mueven en direcciones finitas, el dron recibe recompensas basadas en la calidad de señal experimentada por los usuarios.

Se entrenaron tres agentes con distintas funciones de recompensa: maximización del throughput promedio, uso de umbrales de calidad específicos y el diferencial de throughput promedio entre estados consecutivos. Estos drones se compararon entre sí, con una antena estática central y con un dron ideal que siempre toma decisiones óptimas. Los agentes fueron entrenados usando el algoritmo Deep Q-Network.

Los resultados muestran que los drones entrenados superan la calidad de señal de una estación base estática. Entre ellos, el agente basado en el diferencial de throughput promedio entre estados consecutivos obtuvo el mejor desempeño, acercándose a la calidad del dron ideal.



MODELAMIENTO Y PREDICCIÓN DE VARIABLES RELEVANTES PARA, ANTICIPAR FALLAS EN VENTILADORES DE MINERÍA SUBTERRÁNEA UTILIZANDO REDES NEURONALES

CAMILA GARCÍA-GUAJARDO & GONZALO ACUÑA-LEIVA

Este proyecto aplica modelos NARX con redes LSTM y Transformers para predecir fallas en ventiladores mineros con una semana de antelación. Usando datos reales y metodología CRISP-DM, se comparan arquitecturas mediante métricas y pruebas estadísticas. Los resultados apuntan a mejorar la eficiencia energética y seguridad en minería subterránea.

La ventilación en minería subterránea es fundamental para garantizar la seguridad, salud y condiciones ambientales adecuadas para los trabajadores. Sin embargo, su funcionamiento continuo representa uno de los mayores consumos energéticos en este sector, alcanzando entre un 40% y 50% del total de energía utilizada. Las fallas inesperadas en los ventiladores no solo interrumpen la operación minera, sino que también aumentan los riesgos para el personal y generan costos elevados por mantenimientos correctivos.

Pese a la importancia crítica de estos equipos, existe una brecha en la aplicación de inteligencia artificial, específicamente redes neuronales, para modelar y predecir su comportamiento. La escasa documentación y la falta de estudios comparativos limitan la capacidad de prevenir fallas, lo que plantea una oportunidad para aplicar soluciones innovadoras.

En respuesta a esta necesidad, se propone desarrollar un sistema predictivo utilizando modelos NARX implementados con redes neuronales tipo LSTM y Transformers, con el objetivo de anticipar fallas en ventiladores mineros al menos con siete días de antelación. El proyecto sigue la metodología CRISP-DM, avanzando desde la comprensión del negocio y el estado del arte, hasta la preparación de datos y el desarrollo de modelos.

El conjunto de datos, proporcionado por una faena minera nacional, fue depurado y normalizado para eliminar valores atípicos y dividirse en subconjuntos de entrenamiento, validación y prueba. PÓSTERiormente, se entrenaron modelos NAR y NARX con distintos objetivos de predicción, utilizando el optimizador Adam, validación cruzada y búsqueda aleatoria de hiperparámetros. El desempeño de los modelos se evaluó con métricas como MSE y MAE, y se aplicaron pruebas estadísticas (Friedman y post-hoc de Nemenyi) para determinar diferencias significativas entre arquitecturas.

Hasta la fecha, los modelos basados en LSTM han entregado resultados promisorios. La comparación de arquitecturas permitirá seleccionar la solución más eficaz para su futura aplicación en un proyecto FONDEF. Se espera que el desarrollo de estos modelos predictivos sea un aporte sustancial para mejorar la gestión del mantenimiento, reducir costos y prevenir interrupciones críticas en la operación minera subterránea.



GRIPROUTE: INNOVACIÓN TECNOLÓGICA PARA ESCALADORES Y GIMNASIOS DE BOULDER

IAN RICKMERS-BLAMEY & ARTURO ÁLVAREZ-CEA

GripRoute es una aplicación móvil en desarrollo que permite a los escaladores crear y visualizar rutas de boulder de manera automatizada, mejorando su rendimiento y experiencia. También ofrece herramientas para que gimnasios promocionen sus servicios, organicen competencias y accedan a estadísticas de uso, conectando mejor con sus clientes y aumentando su visibilidad.

La escalada, tanto como deporte como actividad recreativa, ha experimentado un notable crecimiento en popularidad en los últimos años. Sin embargo, a pesar de este auge, persisten necesidades insatisfechas en el mercado, especialmente en cuanto a la capacidad de los escaladores para visualizar y planificar sus rutas antes de escalar. Esta limitación no solo afecta la efectividad en la ejecución de las rutas, sino también el disfrute general de la experiencia de escalada. Asimismo, los gimnasios de escalada enfrentan desafíos para promocionar sus servicios y conectar de manera personalizada con sus clientes, lo que subraya la necesidad de una solución innovadora que aborde ambos problemas de manera integral.

GripRoute surge como respuesta a estas necesidades no cubiertas. Es una aplicación móvil diseñada para permitir a los escaladores crear y visualizar rutas de boulder de manera automatizada antes de intentar escalarlas, sin necesidad de tener conocimientos avanzados de escalada. Esta capacidad de planificación anticipada permite a los usuarios mejorar su rendimiento y optimizar su experiencia en el muro de escalada.

Paralelamente, GripRoute ofrece a los gimnasios herramientas para promocionar sus instalaciones, organizar competencias informales y acceder a estadísticas detalladas de uso. Esto les permite mejorar la experiencia del cliente, aumentar su visibilidad y atraer a más usuarios.

Actualmente, GripRoute está en desarrollo, mediante una metodología que combina investigación de mercado, análisis de la competencia y encuestas tanto a escaladores como a gimnasios. Además, se ha incluido el desarrollo de un prototipo inicial.

Este proyecto no solo aborda una necesidad crítica en la comunidad de escalada, sino que también ofrece una solución escalable que beneficia tanto a escaladores individuales como a gimnasios. Esto posiciona a GripRoute como una herramienta indispensable en el mercado de la escalada, con un gran potencial para transformar la forma en que los escaladores planifican y ejecutan sus rutas, así como para mejorar la visibilidad y conexión de los gimnasios con sus clientes.



ESTIMACIÓN DE PARÁMETROS ESTRUCTURALES DE DISCOS PROTOPLANETARIOS MULTIBANDA UTILIZANDO APRENDIZAJE AUTOMÁTICO

ARIEL ARGOMEDO-MADRID & FERNANDO RANNOU-FUENTES

Este trabajo propone el uso de Vision Transformers para estimar parámetros estructurales de discos protoplanetarios observados por ALMA, reemplazando el método MCMC por un enfoque más eficiente. Se espera lograr estimaciones precisas con un margen de ± 2 desviaciones estándar, optimizando la carga computacional y mejorando el análisis de estos objetos astronómicos.

Un disco protoplanetario es una formación circumestelar compuesta de gas y polvo que gira alrededor de una protoestrella. El estudio de estas estructuras es fundamental para los astrónomos, ya que permite comprender el origen y la evolución de los sistemas planetarios.

Actualmente, el radiotelescopio ALMA (Atacama Large Millimeter/submillimeter Array) facilita la observación de discos protoplanetarios mediante una técnica conocida como radiointerferometría. Esta técnica combina señales de múltiples antenas para obtener imágenes de alta resolución en longitudes de onda milimétricas y submilimétricas. Para determinar los parámetros que describen la morfología del disco, los astrónomos utilizan el método de Markov Chain Monte Carlo (MCMC), que, aunque preciso, es costoso en tiempo y poder de cómputo. En su lugar, este trabajo propone una alternativa basada en Vision Transformers para estimar parámetros estructurales de los discos observados en diferentes frecuencias dentro de la misma banda del continuo, con un margen de error de ± 2 desviaciones estándar.

El enfoque se desarrollará siguiendo el método científico, que incluye la recopilación de datos (síntesis de imágenes utilizando modelos de transferencia radiativa del polvo), el desarrollo o adaptación de un modelo Vision Transformer y su implementación. Finalmente, se realizará un análisis de los resultados obtenidos. Se espera como resultado la estimación de cinco parámetros estructurales que describan la morfología del disco protoplanetario.

El modelo propuesto estará diseñado para recibir como entrada un disco observado en cualquier banda de ALMA (4, 6 u 8) y generar estimaciones precisas de los parámetros estructurales con un margen de error de ± 2 desviaciones estándar respecto a los valores reales.

Es importante señalar que no existen herramientas automatizadas para generar los datos necesarios para entrenar este modelo, lo que convierte este proceso en un desafío que requiere ajustes precisos de parámetros para obtener resultados adecuados. Además, las técnicas actuales para estimar parámetros estructurales, como el MCMC, pueden tardar hasta una semana o más debido a la carga computacional requerida, lo que subraya la necesidad de enfoques más rápidos y eficientes como el planteado en este trabajo.



OPTIMIZACIÓN DEL ENTRELAZAMIENTO CUÁNTICO EN SISTEMAS GRANDES MEDIANTE CÓMPUTO DE ALTO RENDIMIENTO

DANIEL CALDERÓN-RIVAS & PABLO ROMÁN-ASENJO

Este proyecto desarrolla algoritmos HPC para optimizar el entrelazamiento cuántico en sistemas con muchos qubits. Se utilizan estrategias de optimización avanzadas basadas en operadores por bloques y descenso dual. El objetivo es mejorar la eficiencia y escalabilidad de sistemas cuánticos, facilitando su implementación práctica en tecnologías cuánticas de gran escala.

La computación cuántica, basada en principios de la mecánica cuántica como la superposición y el entrelazamiento, ofrece un enfoque revolucionario para resolver problemas complejos. El entrelazamiento cuántico, en particular, es un recurso esencial en el procesamiento de información cuántica, ya que permite generar correlaciones no locales entre qubits. Sin embargo, a medida que aumenta el número de qubits, la gestión del entrelazamiento se vuelve exponencialmente más compleja, dificultando la escalabilidad de algoritmos cuánticos.

Los métodos actuales para cuantificar y optimizar el entrelazamiento no son escalables y resultan ineficientes en sistemas cuánticos de gran tamaño, limitando su aplicabilidad en contextos reales. Este proyecto propone utilizar técnicas de cómputo de alto rendimiento (HPC) para desarrollar algoritmos capaces de optimizar el entrelazamiento en sistemas con un número creciente de qubits, superando los obstáculos actuales de complejidad y eficiencia.

El objetivo principal es diseñar algoritmos HPC que permitan gestionar el entrelazamiento de forma eficiente, reduciendo la carga computacional asociada y mejorando el rendimiento de las tecnologías cuánticas. Para ello, se utilizará el enfoque de unificación de medidas de entrelazamiento propuesto por Modi et al. (2010), formulando el problema como una optimización basada en operadores parametrizados por bloques. Se aplicarán estrategias de descenso dual combinadas con heurísticas para explorar el espacio de estados, junto con técnicas que reduzcan los requerimientos de almacenamiento en memoria.

Los algoritmos desarrollados serán evaluados en simulaciones de sistemas cuánticos grandes, analizando su escalabilidad y eficiencia computacional. Con esto, se espera avanzar hacia sistemas cuánticos más eficientes y viables, permitiendo un desarrollo más robusto de tecnologías cuánticas a gran escala.



EXPLICABILIDAD ORIENTADA A OBJETIVOS EN MÉTODOS DE APRENDIZAJE POR REFUERZO BASADOS EN GRADIENTE DE ESTRATEGIAS

BENJAMÍN LÓPEZ-HUARNEZ & JAVIER BALADRÓN-PEZOA

Este trabajo propone un modelo predictivo que estima las probabilidades de éxito de agentes entrenados con métodos basados en gradientes de estrategias, facilitando explicaciones más claras de sus decisiones. Evaluado en escenarios discretos y continuos, el modelo demostró ser confiable, proporcionando una solución al problema de confianza en sistemas de IA.

Una de las principales preocupaciones respecto al avance actual de la Inteligencia Artificial (IA) es su baja confiabilidad, causada principalmente por la falta de comprensión sobre su funcionamiento interno y la forma en que toma decisiones. Esta desconfianza, además de limitar el uso de la IA, podría influir en su futuro desarrollo y restringir su aplicación en sectores con gran potencial de innovación.

Por naturaleza, la confianza en la interacción humano-computador aumenta con experiencias pasadas. Sin embargo, el alto riesgo percibido al usar IA en ciertos contextos dificulta lograr interacciones efectivas. Una solución para reducir este riesgo es diseñar sistemas capaces de explicar sus decisiones, lo que representa un desafío especialmente complejo en los modelos de caja negra.

Este problema se agrava en el aprendizaje por refuerzo, donde un agente aprende a resolver tareas mediante señales de recompensa. La explicación lógica típica es que se toman decisiones para maximizar la recompensa, lo cual puede resultar poco intuitivo para usuarios sin conocimiento del método de entrenamiento.

Esto se complica aún más con métodos sofisticados, como los basados en gradientes de estrategias (policy gradient), donde es difícil determinar reglas o causas comprensibles para el usuario.

La solución propuesta en este trabajo consiste en desarrollar un modelo predictivo capaz de estimar las probabilidades de éxito de un agente entrenado con métodos basados en gradientes de estrategias. Este modelo utiliza dichas probabilidades para generar explicaciones más claras de las decisiones del agente.

Se implementaron dos escenarios experimentales: uno discreto y otro continuo. En ambos casos, se entrenó a un agente con métodos basados en gradientes de estrategias, y PÓSTERiormente se desarrolló un modelo predictivo diseñado para calcular las probabilidades de éxito del agente en un estado dado y al realizar una acción específica. Diversas pruebas evaluaron la confiabilidad del modelo, validándolo en distintas condiciones para garantizar predicciones consistentes y precisas.

El resultado fue un modelo predictivo que explica decisiones basadas en objetivos de agentes entrenados con gradientes de estrategias. Este modelo, probado en escenarios discretos y continuos, demostró ser confiable y facilitar explicaciones más comprensibles.



FRAMEWORK EVOLUTIVO PARA EL DISEÑO DE ALGORITMOS NEUROMÓRFICOS

BASTIÁN LOYOLA-JARA & JAVIER BALADRÓN-PEZOA

Este trabajo aborda el desarrollo de algoritmos para computación neuromórfica, proponiendo un framework basado en neuro-evolución (NEAT) implementado en ANNarchy. Se prueba en tareas clásicas como XOR y carro de poste, evaluando codificaciones y modelos neuronales. Los resultados muestran rendimiento competitivo frente a métodos tradicionales de aprendizaje por refuerzo.

Las arquitecturas tradicionales basadas en el modelo de von Neumann presentan limitaciones debido a la separación física y funcional entre las unidades de procesamiento y memoria. Aunque este diseño ha sido clave en el desarrollo de la computación moderna, conlleva un alto consumo energético y reduce la eficiencia operativa, especialmente en aplicaciones de inteligencia artificial.

La computación neuromórfica busca superar estas limitaciones imitando el funcionamiento del cerebro humano, con el objetivo de alcanzar su eficiencia energética y capacidad de adaptación. Un chip neuromórfico emula la dinámica de neuronas y sinapsis biológicas, permitiendo procesar información con menor consumo de energía y mayor eficiencia.

Avances recientes en hardware, como los chips Loihi de Intel y el sistema TrueNorth de IBM, han demostrado capacidades destacadas en el procesamiento en tiempo real de modelos neuronales a gran escala. Sin embargo, el diseño de algoritmos efectivos para entrenar y optimizar redes neuromórficas sigue siendo un desafío.

Este trabajo propone un framework evolutivo para la creación de algoritmos neuromórficos, basado en la técnica de neuro-evolución NEAT. El objetivo es mejorar la capacidad y eficiencia de redes neuronales de espigas en entornos dinámicos y aplicaciones en tiempo real. La implementación se realiza en el simulador ANNarchy, compatible con distintas configuraciones de hardware neuromórfico.

Actualmente se estudian dos problemas no lineales: el XOR y el del carro de poste. Para cada uno se analizan distintas codificaciones de información neuromórfica y modelos neuronales, además del efecto de los hiperparámetros de NEAT, optimizados con algoritmos bayesianos mediante la librería Optuna. Los resultados muestran que los algoritmos generados pueden competir con métodos tradicionales de aprendizaje por refuerzo.



SOLUCIÓN NEUROMÓRFICA AL PROBLEMA DE BANDIDOS CONTEXTUAL NO ESTACIONARIO MEDIANTE UN MODELO COMPUTACIONAL DE LOS GANGLIOS BASALES

JOSÉ PINEDA-MOYA & JAVIER BALADRÓN-PEZOA

Algoritmo neuromórfico que resuelve el problema de bandidos contextual no estacionario mediante un modelo de ganglios basales. Este trabajo integra múltiples contextos y el circuito STN-GPe. Aprende rápido, ajusta sinápsis y prioriza acciones previamente recompensadas, superando métodos tradicionales en adaptación y eficiencia.

La computación neuromórfica, inspirada en el funcionamiento del cerebro, difiere de la computación tradicional basada en la arquitectura de Von Neumann. En lugar de utilizar CPU y unidades de memoria separadas, emplea neuronas y sinapsis, lo que introduce desafíos importantes en el diseño de algoritmos. Estos sistemas son especialmente prometedores en aprendizaje automático, ya que implementan de forma natural un estilo de cómputo similar al de las redes neuronales artificiales.

Existen distintos enfoques para desarrollar algoritmos neuromórficos, cada uno con limitaciones propias. Una alternativa novedosa consiste en inspirarse en la neuroanatomía, estudiando y emulando la organización de estructuras cerebrales asociadas al tipo de problema a resolver, para crear modelos matemáticos que puedan ser simulados.

El problema de los bandidos es un ejemplo clásico en aprendizaje por refuerzo. Aunque su versión básica es sencilla, su complejidad aumenta con variantes como la contextual no estacionaria, donde las recompensas cambian con el tiempo y dependen del contexto. Estas variantes tienen aplicaciones en sistemas de recomendación.

En este trabajo se extiende un modelo neurocomputacional de los ganglios basales, conjunto de núcleos cerebrales involucrados en el aprendizaje por refuerzo, para desarrollar un algoritmo neuromórfico capaz de resolver el problema de bandidos contextual no estacionario. En particular, se explora el circuito cerrado STN-GPe, usualmente ignorado en otros modelos.

Se modificó el modelo para incorporar múltiples contextos y así abordar directamente la versión contextual no estacionaria del problema. Cada modificación fue evaluada experimentalmente, analizando sus efectos y límites.

La tarea considerada incluye múltiples bloques, cada uno con su propia acción correcta que cambia según diversas condiciones. En las simulaciones se evaluó el rendimiento del modelo, el sesgo en la exploración tras cambios en la recompensa y los pesos sinápticos del circuito STN-GPe.

Los resultados muestran que el algoritmo neuromórfico es capaz de resolver eficazmente el problema, presentando un sesgo exploratorio hacia acciones previamente recompensadas. Esta propiedad le permite adaptarse más rápido a los cambios del entorno y superar en eficiencia a algoritmos tradicionales.



USO DE TÉCNICAS DE MACHINE LEARNING PARA LA DETECCIÓN DE ANOMALÍAS A PARTIR DE IMÁGENES ASTRONÓMICAS

NICOLÁS SEPÚLVEDA-BALLESTEROS & FERNANDO RANNOU-FUENTES

Se desarrolla un sistema no supervisado de detección de anomalías en imágenes astronómicas tipo stamps mediante redes autoencoder. Utilizando datos del ZTF, el sistema aprende patrones normales para identificar eventos inusuales, facilitando su clasificación en observatorios como el futuro LSST y contribuyendo al estudio de fenómenos astronómicos transitorios.

Para el estudio de fenómenos astronómicos transitorios y de corta duración, es crucial contar con flujos de trabajo que permitan generar alertas en tiempo real ante eventos anómalos, facilitando su detección y seguimiento. Ejemplos de estos fenómenos incluyen estrellas variables, supernovas o galaxias activas (AGN). Actualmente existen proyectos denominados brokers astronómicos, como el chileno ALeRCE, que procesan las alertas generadas por telescopios que realizan barridos sistemáticos del cielo, como ZTF y, próximamente, el LSST, a cargo del Observatorio Vera Rubin en el norte de Chile.

El LSST generará una cantidad masiva de datos a partir de su censo del cielo, lo que plantea el desafío de desarrollar sistemas capaces de filtrar e identificar automáticamente eventos inusuales. En este contexto, se propone construir un sistema de detección de anomalías basado en imágenes astronómicas denominadas stamps. Cada alerta contiene tres tipos de imágenes: science (la imagen más reciente), reference (imagen de referencia con alta relación señal-ruido) y difference (resultado de la resta entre ambas).

Este proyecto plantea el uso de técnicas de machine learning no supervisado para procesar estas imágenes, entrenando redes autoencoder con alertas sin clasificar. La red aprenderá los patrones de comportamiento "normal" en las imágenes, permitiendo detectar anomalías relevantes mediante el error de reconstrucción. Aunque se utilizarán datos del ZTF, la metodología es aplicable a cualquier observatorio que emita alertas en formato stamps. El objetivo es identificar anomalías significativas para su PÓSTERior clasificación y análisis científico.



USO DE POLINOMIOS ORTOGONALES DISCRETOS PARA EL PREPROCESAMIENTO EN LA SÍNTESIS DE IMÁGENES

DAVID VALERO-CROMA & PABLO ROMÁN-ASENJO

Esta tesis propone el uso de polinomios discretos ortogonales para mejorar la síntesis de imágenes en radiointerferometría, reduciendo distorsiones causadas por muestreo irregular. A través de reortogonalización de Gram-Schmidt, se transforman datos a grillas regulares. Se evalúa su rendimiento con simulaciones y datos reales, comparando con técnicas tradicionales como gridding.

La tesis titulada “Utilización de polinomios discretos ortogonales para el preprocesamiento en síntesis de imágenes” de Hugo Francisco Arenas Aros aborda una problemática clave en el procesamiento de datos de radiointerferometría, especialmente relevante en proyectos astronómicos como ALMA. La calidad de las imágenes generadas es fundamental para el análisis científico, ya que artefactos o imperfecciones pueden inducir a interpretaciones erróneas sobre las características de los objetos celestes observados. La principal dificultad radica en el muestreo incompleto de la transformada de Fourier de las señales de radiofrecuencia captadas por los interferómetros, lo que genera irregularidades en las imágenes reconstruidas.

El enfoque propuesto en la tesis utiliza polinomios discretos ortogonales como técnica de preprocesamiento, antes de realizar la síntesis de imagen. A diferencia del método tradicional de gridding, esta propuesta permite extrapolar los datos con mínima varianza, reduciendo distorsiones comunes y mejorando la fidelidad de la reconstrucción.

La técnica emplea la versión moderna de la reortogonalización de Gram-Schmidt (2022), aplicada en dos dimensiones, para generar polinomios complejos discretos ortogonales que transforman los datos irregulares a una grilla regular.

Este método tiene el potencial de mejorar significativamente la calidad de las imágenes, especialmente en escenarios con muestreo incompleto o contaminación por ruido. Si bien el enfoque implica un mayor costo computacional, la mejora en la precisión de las imágenes lo posiciona como una alternativa viable en contextos donde la fidelidad visual es prioritaria.

La tesis plantea extender el estudio a una evaluación amplia, utilizando múltiples conjuntos de datos –tanto simulados como reales– para analizar su rendimiento cualitativo. Las pruebas se realizarán en distintos contextos: variando el tamaño de píxel, el rango de frecuencias, el espaciamiento de las antenas y las polarizaciones. Esto permitirá obtener una comprensión profunda de las ventajas, limitaciones y potenciales aplicaciones del método propuesto.



RECONSTRUCCIÓN DE IMÁGENES INTERFEROMÉTRICAS CON DATOS POLARIZADOS MEDIANTE ALGORITMOS DE REGULARIZED MAXIMUM LIKELIHOOD

LEO VERGARA-SEPÚLVEDA & MIGUEL CÁRCAMO-VÁSQUEZ

Este trabajo presenta un método de reconstrucción de imágenes de polarización en radiointerferometría usando optimización no lineal con el estadístico χ^2 . Comparado con el método CLEAN, estándar actual, el enfoque propuesto genera resultados similares, permitiendo mapear la polarización de fuentes astronómicas con precisión, incluso en presencia de imperfecciones instrumentales.

Un radiointerferómetro es un conjunto de antenas que detecta señales eléctricas de ondas en radiofrecuencia provenientes del cielo. Estas señales son transformadas en datos denominados visibilidades, que permiten generar imágenes de regiones del cielo y, a partir de ellas, estudiar las características físicas de fuentes astronómicas, como su estado de polarización. Sin embargo, debido a distintos procesos físicos e instrumentales, las imágenes obtenidas presentan imperfecciones. Para abordar este problema, existen métodos de síntesis de imágenes que permiten separar el objeto de estudio del fondo ruidoso.

Este trabajo propone un enfoque alternativo para la reconstrucción de imágenes de polarización, basado en métodos de optimización no lineal. En particular, se implementa una técnica de reconstrucción utilizando el estadístico chi-cuadrado (χ^2), con el fin de estimar la distribución espacial de la polarización a partir de las visibilidades. Las imágenes obtenidas mediante esta implementación se comparan con las generadas por el método CLEAN, actualmente el estándar más utilizado en la síntesis de imágenes radioastronómicas. Los resultados muestran que el método propuesto alcanza niveles de reconstrucción similares a CLEAN, logrando generar mapas de polarización precisos para fuentes astronómicas.



MODELO DE DEEP LEARNING PARA ESTIMAR LA DINÁMICA EN DISCOS PROTOPLANETARIOS

ALMA VIDAL-BORMUTH & PABLO ROMÁN-ASENJO

Este trabajo utiliza redes LSTMConv para predecir la dinámica de discos protoplanetarios, estimando el campo de vectores de velocidad a partir de imágenes simuladas generadas con FARGO3D y RADMC3D. Se evalúan diferentes arquitecturas y configuraciones, buscando mayor eficiencia frente a simulaciones tradicionales. El principal desafío es la generación de datos suficientes.

Los discos protoplanetarios son estructuras extensas de gas y polvo donde se lleva a cabo la formación planetaria. Su estudio es fundamental para comprender este proceso y evaluar teorías como la acreción del núcleo y la inestabilidad gravitacional, cuyos fundamentos dependen del movimiento del gas y polvo en estas estructuras. Sin embargo, esta disciplina enfrenta desafíos importantes, como la complejidad inherente de los discos y las limitaciones para observar directamente los procesos en su interior. Aunque las simulaciones actuales ayudan a abordar estas limitaciones, suelen ser ineficientes.

Por ello, este trabajo propone el uso de técnicas de aprendizaje automático para estimar la dinámica de los discos protoplanetarios, específicamente mediante redes LSTMConv (Convolutional Long Short-Term Memory). El objetivo principal es desarrollar y entrenar una red neuronal que prediga el campo de vectores de velocidad en estos discos a partir de una serie de imágenes que representan observaciones en diferentes momentos.

Para el entrenamiento, se generan datos a partir de simulaciones realizadas con FARGO3D y RADMC3D, que permiten crear imágenes (.fits) de discos en distintos periodos de tiempo y obtener sus correspondientes campos de velocidades.

La metodología incluye, una vez generadas las simulaciones, un procedimiento experimental en el que se evalúan al menos tres arquitecturas diferentes de redes LSTMConv bajo diversas configuraciones, como variaciones en parámetros de optimización, tamaño de batch, pasos de tiempo entre imágenes de disco, entre otros. El modelo se entrena con un enfoque supervisado, dividiendo los datos en conjuntos de entrenamiento, validación y prueba para garantizar su capacidad de generalización. El rendimiento se mide comparando las predicciones del campo de vectores de velocidad con los datos simulados, utilizando métricas como el error cuadrático medio y el coeficiente de determinación.

Se espera que este enfoque proporcione una alternativa más eficiente que las simulaciones tradicionales. Los resultados preliminares sugieren que la red neuronal puede captar patrones complejos en el movimiento del gas y polvo, aunque presenta problemas de estancamiento en el aprendizaje debido a la falta de datos, lo que constituye el principal reto de este trabajo.





DEPARTAMENTO DE
**INGENIERÍA
INFORMÁTICA**

[A.R.R.A.Y]

APPLIED RESEARCH DAY

Tercera Jornada de Investigación Aplicada - DIINF 2024

23 de octubre 2024