

Tesis Expuestas



Hemos seleccionados las últimas 4 Tesis de Doctorado y 4 de Maestría que se han defendido, de modo de tener una mirada rápida al conocimiento e innovación generados en las mismas.

Las 4 Tesis de Doctorado abarcan áreas de conocimiento diferentes, lo cual queda reflejado en un rápido análisis del alcance de los resúmenes que se presentan:

- "Propuesta Metodológica y Arquitectónica para el Desarrollo de Juegos Serios Orientados a la Formación de Competencias Tecnológicas"
- "Integración de métodos de descubrimiento de conocimiento embebido en fuentes de información desestructuradas"
- "Métodos y herramientas de desarrollo y evaluación de refactorings para la mejora de la experiencia de usuario en aplicaciones web".
- "Análisis de datos educativos aplicado en el estudio de la incidencia de factores socioeconómicos en el rendimiento escolar"

Las cuatro Tesis de Maestría corresponden a diferentes carreras acreditadas en el Postgrado:

1 Tesis de Maestría en Cómputo de Altas Prestaciones:

"Coplanificación de procesos maleables de aprendizaje automático mediante contenedores"

1 Tesis de Maestría en Ingeniería de Software:

"Soporte al proceso de análisis e interpretación de variantes genéticas en Exoma humano"

1 Tesis de Maestría en Inteligencia de Datos orientada a Big Data:

"Clasificación automática de correos electrónicos"

1 Tesis de Maestría en Tecnología Informática aplicada en Educación:

"Laboratorio virtual de electrónica básica para alumnos universitarios dentro de aula extendida"

Tesis en este número

Doctorado en Ciencias Informáticas

Sandi Delgado Juan Carlos
Rodríguez Juan Manuel
Gardey Juan Cruz
Pincay Ponce Jorge Ivan

Maestría

CÓMPUTO DE ALTAS PRESTACIONES

2023

Libuti Leandro

INGENIERÍA DE SOFTWARE

2023

Moreyra Cecilia

INTELIGENCIA DE DATOS ORIENTADA A BIG DATA

2023

Fernandez Juan Manuel

TECNOLOGÍA INFORMÁTICA APLICADA EN EDUCACIÓN

2023

Navarria Leonardo

Propuesta Metodológica y Arquitectónica para el Desarrollo de Juegos Serios Orientados a la Formación de Competencias Tecnológicas

Tesis completa:

<http://sedici.unlp.edu.ar/handle/10915/154614>



Tesista

Juan Carlos Sandí Delgado

Directora

Dra. Patricia Alejandra Bazán

MOTIVACIÓN

Ante el surgimiento de las de las Tecnologías de la Información y de la Comunicación (TIC) y su rápido avance e impacto a nivel educativo, las Instituciones de Educación Superior (IES) se ven obligadas a enfrentarse al reto de mantenerse actualizadas ante las nuevas demandas de la sociedad actual (Vega-Hernández et al., 2018). Asimismo, (Cabero et al., 2015) señalan que las TIC brindan a las IES el potencial para la creación de escenarios de aprendizaje que permiten el acceso inmediato tanto de información como de recursos tecnológicos que facilitan la formación permanente del estudiantado y el profesorado en tecnologías digitales (Es importante aclarar que por razones culturales se utiliza en Costa Rica la palabra “docente” como sinónimo de “profesor” o “profesorado”. Por ende, en este documento se utilizan ambos términos como sinónimos).

Aguaded & Cabero (2014), afirman que el principal problema que ha manifestado el profesorado es la actualización profesional, ya que requieren de diversas actividades de capacitación relacionadas a potenciar sus habilidades y competencias tecnológicas, mismas requeridas para utilizar e integrar las TIC en los procesos formativos. En investigaciones realizadas por Teo (2009) y Scherer et al., (2015), se afirma que el proceso de enseñanza y aprendizaje apoyados por las TIC pueden ser innovados mediante la incorporación de juegos serios en las metodologías de trabajo. Lo anterior, debido a que los mismos pueden ser utilizados para la adquisición de habilidades, formación de competencias, cambios de actitud, comportamiento y, pueden potenciar la motivación y el rendimiento del aprendizaje, en esta línea, los juegos serios pueden adquirir un rol protagónico ante la posibilidad de convertirse en un medio para el acercamiento del profesorado a las tecnologías digitales (Boyle et al., 2011; Chang et al., 2012).

Sandí & Bazán (2017) realizaron una investigación con el objetivo de plantear retos y desafíos para la ejecución de juegos serios educativos utilizando las posibilidades que brinda el modelo de SaaS en Cloud Computing. Como resultado del trabajo se obtuvo una lista de características funcionales que los juegos serios educativos adquirirían al ser utilizados como software como servicio.

En esta misma línea, Sandí & Sanz (2018) llevaron a cabo una investigación con el propósito de in-

dar el estado del arte de los juegos serios educativos y sus posibilidades para la adquisición de competencias tecnológicas en la formación del profesorado. Como resultado se obtuvo la elaboración de una propuesta de integración de juegos serios para la formación de competencias tecnológicas en el profesorado y para la mejora de sus actitudes frente a estas tecnologías. Además, se elaboró de una propuesta para la evaluación del grado de aceptación de las tecnologías digitales por parte del profesorado, en particular sobre juegos serios.

Si bien existe variedad de investigaciones que hacen referencia a la posibilidad que ofrecen los juegos serios para la formación de habilidades y competencias digitales, no se evidencian juegos serios diseñados exclusivamente para potenciar habilidades y competencias tecnológicas del profesorado a nivel de educación superior. Asimismo, no se ha localizado o identificado en las referencias bibliográficas analizadas, la existencia de metodologías específicas que orienten o sirvan de guía para la creación y diseño de juegos serios orientados a la formación de competencias digitales.

Los resultados obtenidos en las investigaciones realizadas dan pie y son referente teórico para la elaboración de una investigación que permita proponer una metodología y arquitectura de software que guíe los pasos a seguir durante la creación e implementación de juegos serios que vayan a ser utilizados para la formación de habilidades o competencias tecnológicas. En consecuencia, se plantea aplicar la metodología y la arquitectura de software que se proponga para crear y diseñar un prototipo no funcional de juego serio que pueda utilizarse para la formación de habilidades o competencias tecnológicas.

a) **Objetivo general**

- Definir una metodología y arquitectura de software para el diseño de juegos serios que permitan la adquisición o formación de competencias tecnológicas.

b) **Objetivos específicos**

- Estudiar los conceptos básicos vinculados con las TIC, competencias tecnológicas, juegos serios y desarrollo de software.
- Describir antecedentes de metodologías utilizadas en el diseño de software, en particular de juegos serios, así como las arquitecturas de software y de despliegue utilizadas.

- Elaborar una propuesta metodológica y arquitectónica para el diseño de juegos serios que potencien la adquisición de competencias tecnológicas.
- Validar la metodología a través de un panel de expertos o en el desarrollo de un prototipo no funcional de juego serio, de manera que permita analizar el alcance del aporte de esta tesis y las posibles mejoras por realizar.
- Elaborar un informe con los resultados y las conclusiones a las que se arriba.

APORTES DE LA TESIS

El principal aporte que brinda esta investigación radica en una propuesta metodológica y arquitectónica para orientar y guiar el proceso de diseño e implementación de prototipos de juegos serios, específicamente, los que han sido pensados para ser utilizados en la formación de habilidades o competencias tecnológicas.

Otro de los aportes importantes de la tesis consiste en que, a partir de la metodología desarrollada, permite identificar cual es la población meta más apta para utilizar el juego serio en la formación de habilidades o competencias tecnológicas.

LÍNEAS DE I/D FUTURAS

Se trabajará en el diseño e implementación del prototipo funcional del juego serio **SkillNet**, se considerará la posibilidad de que el mismo sea ofrecido como un software de servicio.

Se debe trabajar en la aplicación de la propuesta metodológica **INTEGRA+506** para el diseño de prototipos de juegos serios en diferentes ámbitos educativos, así como en distintas áreas del conocimiento y competencias.

Integración de métodos de descubrimiento de conocimiento embebido en fuentes de información desestructuradas

Tesis completa:

<http://sedici.unlp.edu.ar/handle/10915/156636>



Tesista

Rodríguez Juan Manuel

Director

Rodolfo Bertone

CoDirector

Patricia Pesado

MOTIVACIÓN

El objetivo o motivación de esta tesis fue la de construir una familia de métodos de extracción de conocimiento, tal que, dado un documento como entrada con información no estructurada pero con conocimiento embebido (por ejemplo un texto en lenguaje natural), sean capaces de generar un conjunto de piezas de conocimiento estructurado, por ejemplo: reglas de producción, subgrafos de una red semántica, etcetera.

Objetivos específicos fueron: construir un método de Open IE para idioma español en el estado del arte, armar un marco de referencia para evaluar diversos métodos existentes Y proponer un método novedoso de OpenIE.

APORTES DE LA TESIS

Creación de métodos de Open IE para idioma español.

Se crearon dos métodos de extracción de conocimiento para la Web en idioma español: TP-OIE-ES y ECMes. Se puede concluir que ECMes es, hoy por hoy, un método de extracción de conocimiento para Web en el estado del arte para idioma español y es el que mejor desempeño tiene en los conjuntos de pruebas utilizados.

Creación de un marco de referencia para la evaluación de los métodos de Open IE

Se puede afirmar que se cumplió el objetivo secundario propuesto con la creación de dos conjuntos de datos para la evaluación de métodos de extracción de conocimiento para la Web, uno en inglés: Reuters-103 y otro en español: conjunto Rodríguez. Ambos conjuntos contienen textos en lenguaje natural y para cada oración o párrafo tienen una serie de relaciones semánticas creadas a mano.

Objetivo secundario: creación de un método de Open IE novedoso

Para cumplir con este objetivo se creó primeramente el método TP-OIE, el cual sirvió para poner a prueba la mecánica propuesta. A partir de los prometedores resultados de ATP-OIE, se diseñaron los métodos: TP-OIE-ES y ECMes para idioma español.

LÍNEAS DE I/D FUTURAS

Todo método de extracción de conocimiento para la Web debe ser comparado, en última instancia, con la capacidad humana de comprender y analizar un texto. Estos métodos permiten, de forma automática y mediante algún procesamiento posterior, analizar, resumir o incluso crear mapas de conocimiento de textos en lenguaje natural. Son utilizados cuando el volumen de datos a procesar, el corpus, es demasiado grande para poder ser analizado por personas. Los errores: falta de precisión, falta de exhaustividad, relaciones semánticas poco informativas y manejo de información subjetiva, son todos errores debidos a la misma causa raíz: falta de entendimiento o de comprensión de un texto.

En paralelo a la proliferación de métodos de extracción de conocimiento para la Web, han aparecido algunos modelos puntuales para intentar crear una inteligencia artificial capaz de “comprender” texto en lenguaje natural o bien acercarse lo más posible a este fin. En particular, se trata de modelos de redes neuronales artificiales, preentrenados. Una de ellas es GPT-3 (2020) y la otra es BERT (2019). A pesar lo promisorio de estos modelos, no logran procesar el lenguaje natural al mismo nivel que un ser humano. Incluso uno de los métodos analizados en el documento de tesis: Multi2OIE, utiliza internamente BERT y aun así no es el método con mejor desempeño.

De lograrse el objetivo de construir una IA capaz de comprender un texto en lenguaje natural de la misma forma en la que lo hace una persona, los métodos de extracción de conocimiento para la Web dejarían de tener sentido, ya que su funcionalidad sería sólo una de las muchas tareas factibles para estos sistemas. Por lo cual, los problemas planteados seguirán abiertos hasta que pueda resolverse de forma satisfactoria la comprensión de textos de forma automática.

Métodos y herramientas de desarrollo y evaluación de refactorings para la mejora de la experiencia de usuario en aplicaciones web.

Tesis completa:

<http://sedici.unlp.edu.ar/handle/10915/95679>



Tesista

Juan Cruz Gardey

Directora

Alejandra Garrido

CoDirectores

Gustavo Rossi

Julián Grigera.

MOTIVACIÓN

La experiencia de usuario (UX) es un factor crucial que determina el éxito de las aplicaciones web. Si bien actualmente se reconoce la importancia de proveer una experiencia de usuario óptima, las prácticas de evaluación y mejora de la UX suelen postergarse. Más allá de los recursos necesarios como expertos en UX y usuarios para hacer pruebas, el tiempo requerido por estas prácticas es el principal motivo por el cual resulta difícil integrarlas en los métodos ágiles ampliamente utilizados hoy en día.

Para facilitar la evaluación de la UX, se han desarrollado métodos automáticos para detectar y solucionar problemas de interacción. Uno de los métodos propuestos es la técnica de refactoring de usabilidad, la cual se basa en aplicar pequeñas transformaciones (refactorings) a través scripts ejecutados en el navegador que sin alterar la funcionalidad de la aplicación pretenden mejorar la UX. Esta técnica fue propuesta como una solución automática, en la que los refactorings aparecen como soluciones definitivas sin antes poder validar cómo impactan en los usuarios.

En esta tesis se desarrolla un conjunto de métodos y herramientas para dar soporte en las tareas de evaluación y mejora de la UX. En primer lugar se presenta un método para explorar alternativas de diseño haciendo uso de la técnica de refactoring. Este método se desarrolla en una herramienta llamada UX-Painter, mediante la cual un diseñador puede crear distintas versiones de una aplicación web utilizando los refactorings para generar los cambios de diseño, sin la necesidad de codificarlos. De esta manera, es posible visualizar, validar y evaluar cambios de diseño en la misma aplicación web sin que éstos tengan que ser implementados.

Respecto a la evaluación de los diseños, no solo es importante determinar el impacto de un cambio de diseño en la experiencia del usuario, sino que además resulta fundamental poder comparar distintas alternativas. Sobre todo porque los refactorings en ciertos casos ofrecen más de una solución posible para un mismo problema. Por este motivo, en esta tesis también se desarrolla una métrica denominada esfuerzo de interacción que permite evaluar y comparar diseños alternativos en páginas web. Se trata de un puntaje asignado por expertos en UX que asigna sobre los distintos widgets de una página web, que son aquellos elementos que se ven

modificados por los refactorings. Para calcularlo automáticamente, se desarrollaron modelos de predicción que toman como entrada diferentes logs de interacción de usuario que reciben el nombre de micro-medidas. El hecho que la métrica sea transparente a los usuarios hace que pueda ser utilizada para evaluar diseños con múltiples usuarios en experimentos online como A/B testing o similares. Como prueba de concepto, se implementa UX-Analyzer, una herramienta web que permite calcular y visualizar el esfuerzo de interacción de distintas versiones de una aplicación web.

Finalmente, una vez que se verifica que ciertos refactorings mejoran la UX, deben ser implementados en la aplicación bajo análisis. Para facilitar esta tarea a los desarrolladores, se presenta un mecanismo para generar una implementación preliminar de los refactorings usando como base los frameworks y librerías utilizados actualmente para el desarrollo de interfaces web.

El trabajo incluye distintas validaciones que comprueban la aplicabilidad de los desarrollos propuestos para facilitar la mejora de la UX.

APORTES DE LA TESIS

Un método de End-user Development para explorar variantes de diseño utilizando la técnica de refactoring.

La incorporación de nuevos Client-side web refactorings al catálogo desarrollado en trabajos anteriores. Un mecanismo de adaptación de estilos que permite personalizar ciertas propiedades estéticas de los cambios de diseño impuestos por los refactorings.

Una herramienta denominada UX-Painter que implementa los 3 puntos anteriores.

La métrica llamada esfuerzo de interacción que sirve para evaluar el funcionamiento de los distintos elementos o widgets de la UI que se ven modificados por los refactorings.

Un conjunto de micro-medidas para cada uno de los 6 tipos de widget considerados por la métrica anterior que cuantifican diferentes aspectos de la interacción del usuario.

Modelos de predicción que permiten estimar el esfuerzo de interacción de un widget utilizando como entrada las micro-medidas calculadas automáticamente.

Una herramienta llamada UX-Analyzer que utiliza

los modelos anteriores para calcular y visualizar el esfuerzo de interacción de distintas versiones de una página web.

LÍNEAS DE I/D FUTURAS

Desarrollar nuevos Client-side web refactorings para contar con más opciones para explorar cambios de diseño.

Continuar con el desarrollo del mecanismo para generar automáticamente una implementación preliminar de los refactorings.

Realizar validaciones de las herramientas UX-Painter y UX-Analyzer en un contexto de uso real para determinar empíricamente su aplicabilidad y potenciales mejoras.

Incorporar nuevos tipos de widget a la métrica de esfuerzo de interacción.

Agregar a UX-Analyzer otro tipo de métricas que también contribuyan a la evaluación de la UX.

Análisis de datos educativos aplicado en el estudio de la incidencia de factores socioeconómicos en el rendimiento escolar.

Tesis completa:

<http://sedici.unlp.edu.ar/handle/10915/156471>



Tesista

Jorge Iván Pincay Ponce

Director

Armando Eduardo De Giusti

MOTIVACIÓN

Esta investigación se desarrolló en el campo de la Minería de Datos Educativos. Articula el análisis de datos con el problema multifactorial del rendimiento académico escolar motivado por la incidencia de factores socioeconómicos. Se han aplicado modelos supervisados, no supervisados y un análisis confirmatorio sustentado en exploraciones estadísticas que preceden a los modelos.

Para muchos niños, dejar la escuela es el final de un proceso de desconexión gradual y de reducción de su participación en el currículo formal y la vida social escolar. La deserción desencadena mayormente en la reprobación de materias a causa de un insuficiente rendimiento, que según Tarik y colaboradores (2021), puede suscitar que varios estudiantes adopten comportamientos antisociales o luego situarse en la franja laboral más baja.

Según la UNESCO (2017), las escuelas empiezan a concebir al análisis de datos como oportunidades educativas significativas para reducir la deserción bajo la premisa que cuando se sepa de estudiantes menos favorecidos se tendrá probabilidades de mejorar su estancia educativa.

La minería de datos educativos favorece la búsqueda de relaciones entre los puntajes de los alumnos, exploración de conocimientos sobre el aprendizaje contextualizado, análisis de factores socioeconómicos que lo afectan de modo clave o tendencias de desarrollo a partir del rendimiento y comportamiento actual (McNeish & Wolf, 2020).

El sistema escolar ecuatoriano consta de tres años enfocados en el fortalecimiento cognitivo, afectivo y psicomotriz; y tres años que abordan destrezas y técnicas de estudio. Los datos analizados incluyen promedios quimestrales y parciales, además, datos socioeconómicos como: estructura familiar, ingreso familiar, parentesco del representante, dirección/distancia domiciliaria, cantidad de personas que viven en casa, escolaridad alcanzada por padres y disponibilidad de servicios básicos como acceso a computadoras e internet.

La pandemia del COVID-19 reveló fragilidades en los sistemas educativos, a octubre de 2020, según la UNESCO (2021), un tercio de los estudiantes del mundo no tenían acceso a la educación online debido a la falta de los dispositivos o conectividad, o sus escasas habilidades digitales.

Entonces: ¿Cómo inciden los factores socioeconómicos en el aprovechamiento escolar?

Objetivos

General: Analizar la incidencia de los factores socioeconómicos en el aprovechamiento académico a nivel escolar y de ese modo contribuir a su entendimiento y mejora, mediante la aplicación de modelos de análisis de datos.

Específicos:

1. Reconocer las aplicaciones de análisis de datos en los problemas del contexto educativo escolar.
2. Preparar los datos acordes con la dimensionalidad a un número efectivo de características.
3. Estudiar comparativamente algoritmos de minería de datos en función del proceso de minería de datos que resulte idóneo.
4. Establecer parámetros e hiperparámetros apropiados a los datos y los modelos.
5. Interpretar los resultados del conocimiento descubierto y su eficiencia según métricas pertinentes a los modelos.

APORTES DE LA TESIS

1. Contribución al área de análisis de datos para el mejoramiento educativo, teniendo en cuenta información disponible en Ecuador (área donde se desempeña el Tesista), pero con resultados proyectables en forma general debido al impacto de la socioeconomía en el rendimiento académico.
2. Análisis de datos académico y socioeconómicos escolares.
3. Revisión de calificaciones finales y de avances progresivos para generar recomendaciones tempranas sustentadas en datos que disipen posibles afectaciones al aprovechamiento.
4. Exploración y configuración los datos con criterios de ingeniería, con hipótesis y sus validaciones en tareas de clasificación y predicción, hasta evidenciar métricas favorables que sugieran resultados interpretables y susceptibles de convertirse en recomendaciones para una apropiación social que beneficie a la educación escolar.

LÍNEAS DE I/D FUTURAS

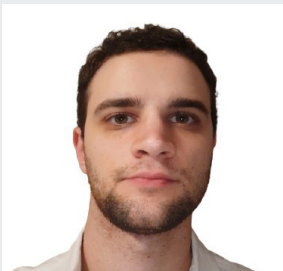
Se aspira a investigar para sustentar acciones respaldadas con datos e Inteligencia Artificial que ayuden a los niños a una estancia escolar más agradable:

- Incorporación de más escuelas de distintas regiones para obtener un abordaje más significativo.
- Revisiones literarias en más bases de datos e idiomas empleando inteligencia artificial para conseguir cantidad y calidad literaria.
- Observación de cada factor socioeconómico y el favorecimiento de medidas proactivas que mitiguen la problemática multifactorial.
- Generación de soluciones de aprendizaje automático interactivas e interpretables con base en normativas emergentes que transparenten a los usuarios la razón de sus decisiones.
- Abordaje de datos psicológicos, estilos de aprendizaje, autoeficacia, ambiente hogareño, motivación u otra información obtenible mediante encuestas periódicas que evidencien su variabilidad situacional temporal. Los psicólogos de las escuelas usualmente recurren a textos (analizables con minería de texto) manuscritos por los alumnos que pueden ser enriquecidos con sus anotaciones para contrastar habilidades motoras finas, lenguaje, memoria y concentración, con su rendimiento.

Coplanificación de procesos maleables de aprendizaje automático mediante contenedores

Tesis completa:

<http://sedici.unlp.edu.ar/handle/10915/153706>



Tesista

Leandro Ariel Libutti

Director

Francisco Igual

CoDirectora:

Laura De Giusti

MOTIVACIÓN

El crecimiento exponencial en interés por el aprendizaje automático en la última década está directamente relacionado con tres avances fundamentales:

1. El desarrollo de mejores algoritmos con aplicaciones directas en muchos campos de la ciencia y la ingeniería;
2. La disponibilidad de cantidades masivas de datos y la viabilidad de almacenarlos y analizarlos de manera eficiente;
3. La aparición de arquitecturas de hardware novedosas, normalmente paralelas y / u homogéneas, que permiten una adecuada explotación de estos algoritmos sobre grandes conjuntos de datos en un tiempo asequible.

La aparición de nuevas arquitecturas y técnicas en HPC ha renovado el interés por el Machine Learning en una gran variedad de problemas, incluyendo aplicaciones de reconocimiento de imágenes, segmentación, reconocimiento de voz, procesamiento de lenguaje natural o traducción de idiomas, entre muchos otros. Tensorflow, Caffe, Keras y PyTorch son frameworks de ML que permiten ocultar detalles de implementación al usuario manteniendo un alto rendimiento tanto en el entrenamiento como en la inferencia de modelos.

Hoy en día, Tensorflow es uno de los frameworks más utilizados. Su diseño se basa en un grafo de ejecución en el cual las operaciones están representadas por nodos y el flujo de datos por tensores (matrices multidimensionales de datos). Para cada operación puede definirse el paralelismo (intra paralelismo). A su vez, se puede elegir la cantidad de operaciones simultáneas que pueden ejecutarse (inter paralelismo). El gran problema que presentan ambos paralelismos es que son definidos antes de la ejecución del algoritmo sin posibilidad de modificarlos en tiempo de ejecución.

La rigidez en el uso de recursos no es en sí misma problemática, siempre que las aplicaciones se ejecuten de forma aislada en el computador. Sin embargo, en entornos en los que múltiples aplicaciones conviven en el mismo computador, y tanto sus puntos de llegada al sistema como los requisitos a nivel de recurso que éstas presentan puedan variar, dicha rigidez puede llevar a situaciones de sobre suscripción, degradando el rendimiento global del sistema y particular de cada aplicación.

Así, parece necesario no solo aplicar técnicas de planificación de procesos y gestión de recursos avanzadas que permitan optimizar el rendimiento global del sistema y particular de las aplicaciones, respetando a la vez los requisitos de calidad de servicio (QoS) sino además hacerlo siendo consciente de la posibilidad de gestionar aplicaciones elásticas.

Por otro lado, la utilización de contenedores como método de virtualización ligera permite una mejor administración de los recursos y portabilidad. Existen múltiples planificadores que permiten aprovechar los beneficios de los contenedores, pero solo permiten llevar a cabo una asignación estática de recursos en el momento de su creación, y en algunos casos reasignación de recursos en tiempo de ejecución; en cualquier caso, las aplicaciones en ejecución dentro del contenedor no se encuentran preparadas para reaccionar ante dicho evento, y por tanto no se adaptarán en ningún caso a la modificación en los recursos asignados al contenedor.

Los objetivos de este trabajo son:

- Modificación del esquema de gestión de recursos dentro de la infraestructura del framework Tensorflow para permitir selección dinámica del paralelismo de las operaciones que conforman el modelo de ML.
- Diseño e implementación de un controlador interno de cada contenedor que permita manejar los recursos computacionales asignados dinámicamente al algoritmo de ML, y de un mecanismo de comunicación entre el sistema y la aplicación confinada en el contenedor.
- Diseño e implementación de un planificador de contenedores que ejecutan algoritmos de ML sobre Tensorflow elástico utilizando técnicas de orquestación que permitan administrar los recursos computacionales del sistema eficientemente.

APORTES DE LA TESIS

En este trabajo se propone el diseño e implementación de un mecanismo completo de elasticidad en el uso de recursos computacionales en el framework Tensorflow, permitiendo la reasignación dinámica de núcleos de cómputo durante la ejecución del algoritmo de ML. Además, se extiende el uso de la elasticidad a contenedores con la implementación de un controlador/cliente que permita administrar los recursos computacionales asignados a los algoritmos de ML que se eje-

cutan internamente. Por último, se implementa un planificador de contenedores elásticos con el fin de gestionar dinámicamente los recursos del sistema entre todos los contenedores activos y definir políticas de planificación que favorezcan el rendimiento global del sistema.

LÍNEAS DE I/D FUTURAS

- Incorporación de nuevas políticas de planificación de contenedores que permitan mejorar el rendimiento alcanzado para las implementadas anteriormente.
- Estudio y análisis de la eficiencia energética del coplanificador.
- Incorporar en el planificador soporte de sistemas heterogéneos utilizando aceleradores tales como GPU, TPU, entre otros.
- Analizar el uso de funciones desarrolladas por proveedores de GPU, como Multi-Stream, Multi-Process Service (MPS), Multi-Instance GPU (MIG) y GPUs virtuales (vCS) para soporte de planificación en tiempo de ejecución y administración de recursos.
- Explorar la administración de recursos en otros frameworks de ML tales como PyTorch y Caffe, con el fin de extender la elasticidad a más aplicaciones y realizar estudios comparativos entre ellos.

Soporte al proceso de análisis e interpretación de variantes genéticas en Exoma humano

Tesis completa:

<http://sedici.unlp.edu.ar/handle/10915/156057>



Tesista

Cecilia Moreyra

Director

Andrés Rodríguez

CoDirector

Alejandra Garrido

MOTIVACIÓN

El propósito principal de esta tesis es dar soporte al médico genetista en el proceso de diagnóstico genético, en particular el acceso a la información sobre la composición genética de sus pacientes, la identificación de los posibles escenarios de detección y diagnóstico y el registro de cada uno de los pasos y datos llevados a cabo en el proceso.

Objetivos específicos

- Modelar el proceso de toma de decisiones para el filtrado y selección de variantes de un exoma relevantes llevado a cabo por el médico genetista para alcanzar un informe de diagnóstico de un paciente.
- Proveer herramientas de asistencia, basadas en una visualización eficaz, para la ejecución de ese proceso de análisis e interpretación con mayor rapidez y sensibilidad.
- Acelerar los tiempos de interpretación.
- Aumentar la sensibilidad (en este caso la tasa de detección de las variantes patogénicas causales de la enfermedad o que sean relevantes para su salud).

APORTES DE LA TESIS

Esta tesis aporta a la comunidad médica, dedicada al área genética, elementos para poder realizar un diagnóstico integral en base al análisis del Exoma Humano y generar una base de datos genéticos propia con el registro del trabajo realizado, pudiendo de esta manera mejorar los tiempos de evaluación diagnóstica, tener la información centralizada y contar con información para diagnósticos futuros.

LÍNEAS DE I/D FUTURAS

Como principal trabajo futuro se sugiere que, en caso de ser necesario porque el volumen de trabajo lo justifique, sea integrada la consulta manual a las bases de datos de datos genéticas tales como OMIM, Phenomizer, entre otras. Podría también construirse un sistema de recomendación basado en la lectura previa de las mencionadas apis.

Podría también a futuro ser desarrollado un módulo de contribución e integración a las bases de datos de estadísticas de enfermedades detectadas en la población argentina, a efectos de tomar decisiones de prevención en las patologías más frecuentes. El diseño de la interfaz de usuario puede utilizarse como punto de partida para un proceso de desarrollo e implementación de la herramienta Exoma Browser.

Clasificación automática de correos electrónicos

Tesis completa:

<http://sedici.unlp.edu.ar/handle/10915/120066>



Tesista

Juan Manuel Fernández

Director

Marcelo Errecalde

CoDirector

Mario Oloriz

MOTIVACIÓN

El objetivo general de este trabajo consiste en estudiar y analizar el conocimiento existente sobre técnicas aprendizaje automático aplicadas a la clasificación automática de textos, particularmente de correos electrónicos, y generar un modelo que aborde un problema concreto. Esto trae aparejados los siguientes objetivos específicos:

- Analizar, describir y sistematizar el estado del arte de la clasificación automática de correos electrónicos.
- Diseñar un proceso general para el tratamiento y clasificación automática de correos electrónicos, intentando categorizar esta problemática dentro de la disciplina general de Minería de Textos, la cual abarca las características y particularidades que se originan en esta forma de comunicación.
- Abordar un estudio experimental a partir del procesamiento de las consultas que los estudiantes de la Universidad Nacional de Luján formulan, mediante correo electrónico, asegurando la calidad de los datos y etiquetando, a partir de especialistas

REQUISITOS

- Idioma: español e inglés (traducir cada uno de los ítems)
- Fotografía personal: Adjuntar al correo una fotografía a color, en formato .jpg, sin marco y de tamaño cuadrado. En lo posible con fondo liso)
- Extensión: 1050 palabras máximo en total (esto incluye ítems Motivación, Aportes de la Tesis, Líneas de I/D futuras)

en el dominio, un subconjunto de los mismos con las temáticas a las que corresponden en ese dominio.

- Consolidar, a partir de lo anterior, una base de conocimiento con los correos electrónicos etiquetados, representados a partir de diferentes estrategias de representación de textos incorporando a los mismos un proceso de curado y generación de atributos estáticos.
- Entrenar un modelo para la clasificación automática de estos correos electrónicos, abordando de manera concreta esta problemática y dimensionando el problema.
- Indagar en el estado del arte de la clasificación semi-supervisada de documentos, proponiendo una estrategia de esta índole para mejorar la performance de clasificación del modelo entrenado.

APORTES DE LA TESIS

En este trabajo se aborda el desafío de analizar, describir y sistematizar el estado del arte del email mining, puntualmente el área que aborda la clasificación automática de correos electrónicos.

Además, se abordó la construcción de un clasificador automático sobre un caso de estudio real que incluyó 1000 correos de entrenamiento divididos en 16 clases. Este conjunto de datos está conformado con las consultas que realizan los estudiantes al staff administrativo de la Universidad Nacional de Luján sobre temas de índole académica. Se procesó y curó una base con 20.000 correos electrónicos y sus respectivas respuestas, asegurando la calidad de los datos y etiquetando, a partir de especialistas en el dominio, un subconjunto de 1000 correos con las temáticas que abordan.

En este sentido, se entrenó un modelo para la clasificación automática de estos correos electrónicos, abordando de manera concreta esta problemática, alcanzando un accuracy del 86% en el mejor de los casos.

Además, con el objetivo de buscar alternativas para mejorar los modelos entrenados, se indagó en el estado del arte de la clasificación semi-supervisada de documentos, logrando, como una de las principales contribuciones de este trabajo, proponer una nueva estrategia para mejorar la performance de clasificación de los modelos entrenados. La hipótesis acuñada para la definición de esta estrategia es que la selección de características representativas de cada clase, combinada con enfoques de recuperación de información, constituyen un método semisupervisado válido y sencillo para la clasificación automática de correos electrónicos. Desde allí, se generó un proceso para la clasificación semi-supervisada de correos electrónicos a partir de la identificación de características claves de cada clase, utilizando tres técnicas de selección de características como la regresión logística, TF-IDF y SS3 y la posterior recuperación de correos etiquetados automáticamente a partir de un enfoque de recuperación de información con un motor de búsqueda de propósito general como Elasticsearch.

Además, se pudo verificar que las estrategias propuestas para el etiquetado semi-supervisado permiten identificar los términos más representativos de cada correo electrónico, al mismo

tiempo que es posible utilizar las ponderaciones definidas por cada técnica para valorar la representatividad de esos términos para cada clase.

Asimismo, se demostró para el conjunto de datos con que se trabajó que estas técnicas de selección de características, utilizadas como estrategias de etiquetado automático a partir de un enfoque de recuperación de información para la clasificación semi-supervisada, mejoran la capacidad de los clasificadores cuando se incorporan las instancias etiquetadas automáticamente a las etiquetadas de forma manual para entrenar el modelo, alcanzando mejoras de entre el 2% y el 6% para muchas de las estrategias abordadas en el marco de esta investigación.

Por último, se presentó, como apéndice, nuevamente la estrategia como una estrategia de oversampling para aplicar al aprendizaje automático en ambientes de datos desbalanceados. En este sentido, nuevamente se demostró, al menos para los datos con que se evaluó, que esta nueva estrategia presentada es competitiva en relación a las estrategias de remuestreo para el balanceo de clases, tanto de undersampling como de oversampling, arrojando valores más altos para las métricas de selección de modelos utilizadas y el conjunto de datos sobre el que se realizaron los experimentos.

LÍNEAS DE I/D FUTURAS

Existen diversas líneas de investigación que quedan abiertas a partir de los resultados de este trabajo:

- Los correos electrónicos en general, y este conjunto de datos en particular, tienen características que no benefician a los modelos de clasificación automática dada la informalidad y los errores de sintaxis recurrentes propios de la dinámica de este medio de comunicación, por lo cual se espera que en conjuntos de datos con textos más depurados se obtengan modelos que logren mejores resultados, para lo cual sería interesante incorporar un proceso de corrección ortográfica sobre el corpus de documentos para la depuración del lenguaje.

- En este trabajo se propone un enfoque semi-supervisado para el etiquetado y clasificación de documentos y se demuestra que el mismo es un enfoque válido y sencillo, al menos para esta colección de documentos. Queda como desafío futuro abordar ajustes en este proceso sobre los

parámetros involucrados, ya sea en términos de las técnicas de selección de características como el ajuste fino del $\$N\$$ con la cantidad de términos representativos para cada clase y el impacto de su calibración en la eficiencia y eficacia de esta estrategia.

- A su vez, para la consolidación de la estrategia de aprendizaje semi-supervisado propuesta, quedan pendientes nuevas pruebas empíricas en otros conjuntos de datos que permitan ratificar la utilidad de la misma, al mismo tiempo que ayude a identificar los contextos más propicios para su utilización.

- Asimismo, en este trabajo se propone enfoque anterior también como estrategia de oversampling para la clasificación automática en un ambiente de datos con clases desbalanceadas. Si bien en el presente estudio se ha limitado la experimentación al dominio de la clasificación automática de correos electrónicos, creemos que la estrategia propuesta es generalizable a otros dominios donde existan disponibles documentos de textos sin etiquetar y, como trabajo futuro, se propone realizar un nuevo trabajo aplicado a un contexto más general de clasificación automática de textos.

- Por otro lado, se plantea la posibilidad de avanzar en la implementación de un modelo de clasificación automática en un entorno de aplicación real, con el fin de realizar una transferencia concreta para facilitar gestión de las respuestas a las consultas en el contexto de la Universidad Nacional de Luján, así como también, analizar la posibilidad de adaptar esta solución -al menos en términos del proceso de construcción propuesto- a otras problemáticas de índole similar.



TESIS Y TESISISTAS



POSGRADOS ACREDITADOS EN INFORMÁTICA



ENCONTRÁ NUESTRAS
ÚLTIMAS PUBLICACIONES EN
<https://postgrado.info.unlp.edu.ar/>

Laboratorio virtual de electrónica básica para alumnos universitarios dentro de aula extendida

Comprobaciones de las leyes básicas
de circuitos eléctricos y aplicaciones
sobre semiconductores

Tesis completa:

<http://sedici.unlp.edu.ar/handle/10915/154689>



Tesista

Leonardo José Navarría

Director

Alejandro Héctor González

CoDirector

María Alejandra Zangara

MOTIVACIÓN

La formación práctica ocupa un lugar relevante dentro de las exigencias docentes, sobre todo con una importancia marcada en las carreras tecnológicas. La metodología habitual necesita la presencia física del alumno en el laboratorio para poder realizar las prácticas y el número de sesiones de prácticas libres está limitado por el tiempo de permanencia en el laboratorio. Debe sumarse a eso el inconveniente que presenta el ejercicio de la docencia en prácticas en laboratorios con equipamiento específico es que el trabajo previo y la preparación del ensayo que el alumno puede hacer fuera del laboratorio se ve limitado al número de horas que el laboratorio puede estar abierto.

La esencia de este proyecto es analizar los conocimientos que aportará el uso de un laboratorio virtual a la hora de realizar prácticas de estudios de circuitos eléctricos. La práctica de laboratorio es una potente estrategia pedagógica para la construcción de competencias procedimentales y por este motivo es utilizada en una gran variedad de programas académicos, usualmente sincronizada con su asignatura teórica correspondiente. El trabajo de Infante Jiménez aborda la importancia los laboratorios presenciales en las carreras de ciencias aplicadas por ejemplo Ingeniería, Físicoquímica, Química, etc., sin embargo, a medida que los modelos educativos se han vuelto más flexibles y enfocados a competencias, la inclusión de las tic ha cambiado radicalmente el concepto de espacio físico; esto ha hecho patente una serie de limitaciones pues a pesar de la enorme importancia que éste tiene para el aprendizaje, no puede ofrecer la versatilidad idónea que se necesita en la actualidad. También es un hecho que la práctica de laboratorio presenta elevados tiempos de respuesta, los cuales tienden a bajar su productividad.

Se considera que el laboratorio remoto dentro del aula virtual "ha sido un recurso importante para posibilitar la enseñanza técnica en la modalidad e-learning ya que ha posibilitado la realización de prácticas, a la vez que mantenía las características esenciales de la modalidad a distancia en cuanto a flexibilidad de horarios y espacio" del trabajo de (Albert Gómez et al) como conclusión del trabajo. Tal arribo de conclusión nuevamente afirma que los laboratorios remotos y virtuales son una excelente herramienta para la construcción del aprendizaje.

La utilización de herramientas de laboratorios virtuales se considera que generará una interacción alumno tutor a través de un entorno virtual formal de enseñanza. La tarea del tutor consistirá en orientar y reorientar los procesos de comprensión y de mediar pedagógicamente a través de las actividades auto estructuradas y progresivas del estudiante a distancia.

El objetivo general de este trabajo es evaluar la efectividad de realizar prácticas de electrónica básica con laboratorios virtuales dentro de un aula extendida.

Con el uso de metodologías de trabajo virtuales se puede lograr que un gran número de alumnos tenga acceso a los contenidos prácticos para poder familiarizarse con instrumentos y componentes de electrónica básica.

En el tema del trabajo se realizan prácticas de circuitos eléctricos de forma virtual, para posteriormente realizarlas dentro de un laboratorio físico dónde se llevarán adelante mediciones con componentes eléctricos. Un aula extendida virtual puede aportar muchas ventajas en la docencia de asignaturas técnicas, pero también resulta de importancia evaluar el impacto que tiene sobre el alumnado.

Los objetivos particulares son:

- Ampliar el alcance de los trabajos prácticos propuestos en las clases de explicación práctica
- Profundizar el análisis de los problemas planteados, llevando las experiencias a mediciones de situaciones reales.
- Analizar los indicadores de expectativas, estimulación, utilidad del laboratorio, uso del laboratorio remoto.
- Proponer orientaciones de uso del laboratorio virtual en el aula de electrónica

APORTES DE LA TESIS

Se realizó una serie de laboratorios virtuales en función a los trabajos prácticos de las asignaturas.

- Se promovió el uso de simuladores como herramienta de aprendizaje.
- Se estableció que las prácticas de laboratorio favorecen los procesos cognitivos.
- Se fomentó la comprobación de resultados al solucionar problemas utilizando simuladores.
- Se fomentó el uso de simuladores para la realización de trabajos en menor tiempo por parte de los alumnos.

• Se trabajó con herramientas de diseño multimedia, como ser diseño web, grabación de videos, edición de fotos.

• Se realizaron las experiencias de laboratorio verificando que los desarrollos sean acordes a los niveles de expertos necesarios.

• Se integraron los laboratorios virtuales en las distintas asignaturas que poseen contenidos de análisis de circuitos eléctricos.

• Se compartieron los contenidos desarrollados en cátedras Electrotecnia y Electrónica y Dispositivos Electrónicos, ambas dos correspondientes a la carrera Ingeniería en Computación (UNLP).

• Se modelaron situaciones reales, al hacer posible el logro de determinados objetivos educativos en cursos en los que se puedan aplicar.

• Se detectó la motivación por parte de los estudiantes ante este tipo de material.

• Se favoreció el proceso de evaluación, ya que los simuladores ofrecen una constante retroalimentación sobre las actuaciones de los alumnos, corrigiendo de forma inmediata los posibles errores.

• Se promovió la innovación con nuevas tecnologías.

LÍNEAS DE I/D FUTURAS

• Profundizar la línea investigativa en relación con las competencias tecnológicas a fin de adaptar el aprendizaje y la evaluación al nuevo paradigma de la formación orientada a las competencias.

• Articular entre los trabajos prácticos presentados en la asignatura con los contenidos de los laboratorios presentados dónde se abordan situaciones reales.

• Mejorar la calidad gráfica de algunos diseños de los laboratorios