



ANEXO 1

FORMATO PARA LA PRESENTACIÓN DE PROYECTOS DE INVESTIGACIÓN CON EL FINANCIAMIENTO DEL FEDU

1. Título del proyecto

Detección de anomalías en electrocardiogramas con autoencoders

2. Área de Investigación

Área de investigación	Línea de Investigación	Disciplina OCDE
Sistemas, Computación e Informática	Inteligencia Artificial y Sistemas Bio-inspirados	Sistemas Cognitivos

3. Duración del proyecto (meses)

12

4. Tipo de proyecto

Individual	<input checked="" type="radio"/>
Multidisciplinario	<input type="radio"/>
Director de tesis pregrado	<input type="radio"/>

4. Datos de los integrantes del proyecto

Apellidos y Nombres	TAPIA CATACORA PABLO CESAR
Escuela Profesional	INGENIERÍA DE SISTEMAS
Celular	943921829
Correo Electrónico	pctapia@unap.edu.pe

I. Título

Detección de anomalías en electrocardiogramas con autoencoders

II. Resumen del Proyecto de Tesis

Un autoencoder es un tipo de red neuronal que busca generar nuevos datos, primero comprimiendo la entrada en un espacio de variables latentes y reconstruyéndolo en la salida en base a la información adquirida. En base a ello, la presente investigación busca detectar las anomalías presentes en un conjunto de electrocardiogramas registrados.

El electrocardiograma (ECG) es un indicador eficaz y no invasivo para la detección y prevención de arritmias. En escenarios del mundo real, las señales de ECG son propensas a estar contaminadas con varios ruidos, lo que puede conducir a una interpretación incorrecta. Por lo tanto, se ha prestado mucha atención a la eliminación de ruido del ECG para un diagnóstico y análisis precisos. Se puede aplicar un codificador automático de eliminación de ruido (DAE) para reconstruir los datos limpios a partir de su versión ruidosa.



En este documento, se propone un DAE que utiliza la fully convolucional network (FCN) para eliminar el ruido de la señal de ECG. Mientras tanto, el DAE basado en FCN propuesto puede realizar la compresión con respecto a la arquitectura DAE. El enfoque propuesto se aplica a las señales de ECG de la base de datos de arritmias del MIT-BIH y las señales de ruido añadidas se obtienen de la base de datos de la prueba de estrés por ruido del MIT-BIH.

Los resultados del estudio serán demostrar la superioridad de FCN en la eliminación de ruido, con root-mean-square error (RMSE) y percentage-root-mean-square (PRD) más bajos, así como una signal-to-noise ratio SNR_{imp} más alta. De acuerdo con los resultados, creemos que el DAE basado en FCN propuesto tiene una buena perspectiva de aplicación en la práctica clínica.

III. Palabras claves (Keywords)

Autoencoders, electrocardiograma, red neuronal, aprendizaje automático

IV. Justificación del proyecto

Actualmente las empresas tienen un ambiente muy competitivo. Las organizaciones modernas buscan optimización de sus recursos, ingresar a nuevos mercados y esto les exige mejorar su funcionamiento con el propósito de lograr una ventaja competitiva sobre sus contrincantes. La toma de decisiones que se lleva a cabo dentro de las organizaciones debe ser rápida, oportuna, fundamentada en información concreta, que permita tomar decisiones eficientes, efectivas y con un bajo costo para la empresa pues de ello dependerá el éxito o fracaso de una organización, (Effy Oz, 2001).

Con el nacimiento de la inteligencia artificial (IA) se plantea la posibilidad de diseñar máquinas "inteligentes", sistemas expertos, algoritmos genéticos, computación molecular o redes neuronales artificiales (RNA). Se define la IA como el estudio de, agentes inteligentes que reciben las percepciones del entorno y llevan a cabo las acciones; también trata las diferentes formas de representar esas funciones, tales como sistemas de producción, agentes reactivos, planificadores condicionales en tiempo real, redes neuronales y sistemas teóricos para las decisiones. En algunas de estas ramas los resultados técnicos van muy por encima de las realizaciones prácticas.

Existen diversas tecnologías de información para la toma de decisiones las cuales presentan un panorama amplio del proceso de toma de decisiones no estructuradas y son las que tienen que ver con la inteligencia artificial, redes neuronales artificiales, Campo de la medicina, sistemas expertos (SE), Sistemas de Apoyo para la Toma de Decisiones (DSS por sus siglas en inglés), Sistemas de Apoyo para la Toma de Decisiones en Grupo (GDSS por sus siglas en inglés) y Sistemas de Información para Ejecutivos (EIS por sus siglas en inglés), entre otros.

V. Antecedentes del proyecto (Incluya el estado actual del conocimiento en el ámbito nacional e internacional. La revisión bibliográfica debe incluir en lo posible artículos científicos actuales, para evidenciar el conocimiento existente y el aporte de la Tesis propuesta. Esto es importante para el futuro artículo que resultará como producto de este trabajo)

Según Russell [2004:663], combina la teoría de la utilidad con la de la probabilidad para así construir un agente inteligente basado en la teoría de la decisión. Esta clase de agente puede adoptar decisiones en situaciones en las que un agente lógico no tiene forma de decidir debido a la presencia de la incertidumbre y objetivos contradictorios. Un agente basado en objetivos maneja una dicotomía entre estados buenos (cumplimiento de objetivos) y malos (no cumplimiento), mientras que un agente basado en la teoría de la

decisión maneja una medida cuantitativa de la calidad de los estados. Una red neuronal artificial es un algoritmo matemático que simula en forma sencilla el funcionamiento en que las neuronas interactúan para procesar datos y aprender con la experiencia, [McLeod, 2000].

El autor Schalkoff (1997:2), en su libro “Artificial Neural Networks”; afirma que, “la estructura de una red está compuesta de un número de unidades interconectadas (neuronas artificiales). Cada unidad tiene una entrada/salida característica y se implementa con un cálculo local o función. La salida de cualquier unidad está determinada por su entrada/salida característica, su interconexión con otras unidades, y posibles entradas externas. La red desarrolla su funcionalidad a través de una o más formas de entrenamiento”.

VI. Hipótesis del trabajo

Las redes neuronales de tipo autoencoders ayudan a identificar las anomalías presentes en datos electrocardiográficos con una precisión alta.

VII. Objetivo general

Analizar la prevalencia de hallazgos electrocardiográficos anormales y su significado clínico

VIII. Objetivos específicos

- Procesar los datos electrocardiográficos utilizando scikit-learn y tensorflow
- Proponer un modelo de autoencoder para detectar anomalías en los sets de datos electrocardiográficos.
- Evaluar el modelo propuesto, con datos electrocardiográficos reales.

IX. Metodología de investigación

En el campo de la investigación se ha demostrado que denoising autoencoder (DAE) es poderoso en el aprendizaje de representaciones de baja dimensión y puede usarse para recuperar entradas con ruido. Aquí, proponemos un DAE basado en fully convolutional network FCN para eliminar los ruidos de las señales de electrocardiogram ECG que contengan alguna alteración.

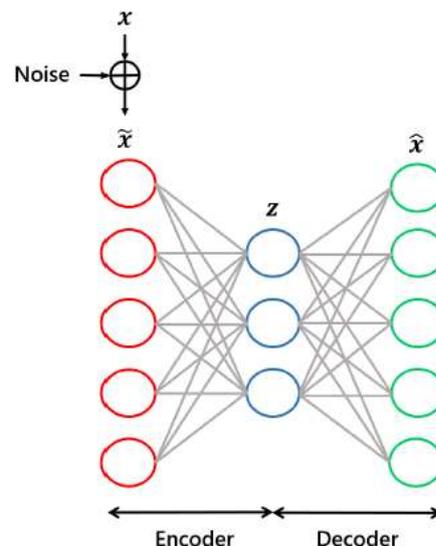


Figura 1. Modelo de una red neuronal autoencoders

La Fully Convolutional Network es un tipo especial de Convolutional Neural Network. CNN normalmente consta de capas convolucionales, funciones de activación, capas de

agrupación máxima y una capa completamente conectada. Las capas convolucionales consisten en un conjunto de filtros que pueden extraer mapas de características para describir las características de los datos de entrada. Diferentes mapas de características en una capa usan diferentes parámetros de filtros con un mapa de características que comparte los mismos parámetros.

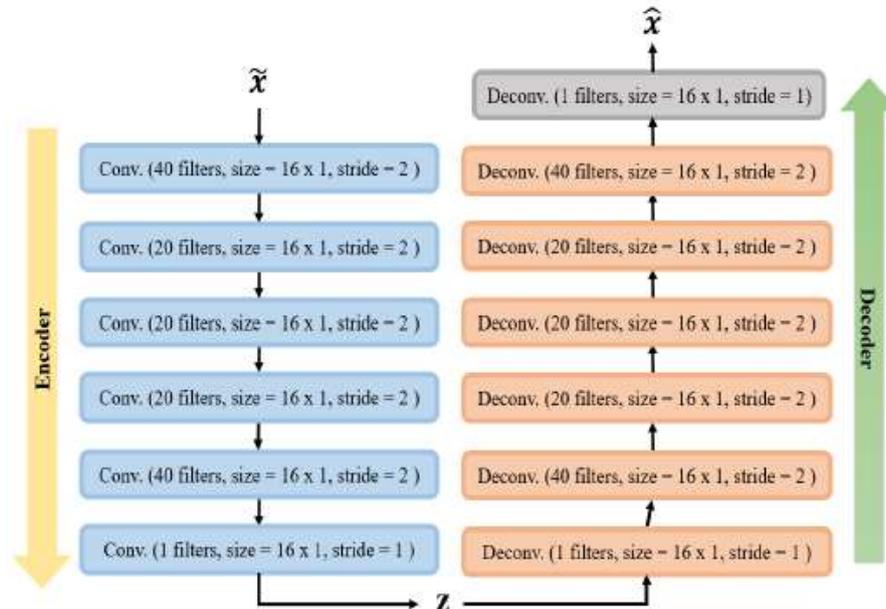


Figura 2. Arquitectura del modelo propuesto de autoencoder

En comparación con la capa totalmente conectada donde cada neurona está conectada a todas las salidas de la capa anterior, las capas convolucionales reducen en gran medida la cantidad de parámetros al compartir parámetros con filtros. La agrupación se utiliza para lograr la invariancia de traslación y rotación, y una operación de agrupación típica es la max-pooling. Las capas de max-pooling realizan un proceso de muestreo descendente al tomar el valor máximo de un cierto alcance del espacio de mapeo, lo que lleva a una dimensión reducida. La capa completamente conectada tiene como objetivo realizar la regresión o la clasificación.

X. Referencias

- Bank, D., Koenigstein, N., & Giryas, R. (2020). Autoencoders. arXiv preprint arXiv:2003.05991.
- Liang, D., Krishnan, R. G., Hoffman, M. D., & Jebara, T. (2018, April). Variational autoencoders for collaborative filtering. In Proceedings of the 2018 world wide web conference (pp. 689-698).
- Cremer, C., Li, X., & Duvenaud, D. (2018, July). Inference suboptimality in variational autoencoders. In International Conference on Machine Learning (pp. 1078-1086). PMLR.
- Kim, Y., Wiseman, S., Miller, A., Sontag, D., & Rush, A. (2018, July). Semi-amortized variational autoencoders. In International Conference on Machine Learning (pp. 2678-2687). PMLR.
- Moor, M., Horn, M., Rieck, B., & Borgwardt, K. (2020, November). Topological autoencoders. In International conference on machine learning (pp. 7045-7054). PMLR.



- Yildirim, O., San Tan, R., & Acharya, U. R. (2018). An efficient compression of ECG signals using deep convolutional autoencoders. *Cognitive Systems Research*, 52, 198-211.
- Chiang, H. T., Hsieh, Y. Y., Fu, S. W., Hung, K. H., Tsao, Y., & Chien, S. Y. (2019). Noise reduction in ECG signals using fully convolutional denoising autoencoders. *IEEE Access*, 7, 60806-60813.
- Kuznetsov, V. V., Moskalenko, V. A., & Zolotykh, N. Y. (2020). Electrocardiogram generation and feature extraction using a variational autoencoder. *arXiv preprint arXiv:2002.00254*.

XI. Uso de los resultados y contribuciones del proyecto

Los resultados se utilizarán en el campo de la educación universitaria, con ello fortalecer el uso e importancia de la inteligencia artificial, las redes neuronales y los modelos en el campo de la medicina humana, específicamente en el campo de la función cardio-vascular.

XII. Impactos esperados

i. Impactos en Ciencia y Tecnología

Implementación de modelos computacionales para ayudar a los profesionales del área médica, hace que se masifique el uso de la tecnología, no solo en el campo de la ingeniería de sistemas, más bien, su uso transversal en las diferentes áreas profesionales.

ii. Impactos económicos

Masificar la accesibilidad a las personas a consultas previas y preventivas, en relación al funcionamiento cardiovascular. Y evitar costos de consultas médicas innecesarias.

iii. Impactos sociales

Valorar el uso de la tecnología para determinar tempranamente las anomalías que pueda presentar los pacientes.

iv. Impactos ambientales

Toda aplicación, uso de la tecnología computacional e inteligencia artificial, no tiene impactos negativos sobre el medio ambiente.

XIII. Recursos necesarios (Infraestructura, equipos y principales tecnologías en uso relacionadas con la temática del proyecto, señale medios y recursos para realizar el proyecto)

Computador/Laptop, Tensorflow, Python, Datos de electrocardiogramas

XIV. Localización del proyecto (indicar donde se llevará a cabo el proyecto)

Puno - Perú

XV. Cronograma de actividades

Actividad	Trimestres
-----------	------------



	E	F	M	A	M	J	J	A	S	O	N	D
Revisión de la literatura	X	X										
Recopilación de datos		X	X									
Pre-procesamiento de datos			X									
Elaboración del 1er informe												
Creación de la arquitectura del modelo de autoencoder				X	X	X	X	X	X			
Elaboración del 2do informe						X						
Entrenamiento del modelo								X	X	X		
Elaboración del 3er informe									X			
Presentación de los resultados										X	X	X
Evaluación y prueba del modelo											X	X
Presentación del informe final de la investigación												X

XVI. Presupuesto

Descripción	Unidad de medida	Costo Unitario (S/.)	Cantidad	Costo total (S/.)
Laptop	Unidad	6400.00	1	6400.00
Electrocardiogramas	Registro	2.00	1200	2400.00
Python	Software	0.00	1	0.00
Impresora	Unidad	800.00	1	800.00
Papel	Millar	25.00	1	50.00
Asesoría profesional	Profesional	2500.00	2	5000.00
Libros	Libro	100.00	15	1500.00
Artículos científicos	Artículo	150.00	20	3000.00
No previstos	Unidad	5000.00	1	5000.00