



UNIVERSIDAD CÉSAR VALLEJO

**ESCUELA DE POSGRADO
PROGRAMA ACADÉMICO DE MAESTRÍA EN
INGENIERÍA DE SISTEMAS CON MENCIÓN EN
TECNOLOGÍAS DE LA INFORMACIÓN**

**Machine Learning para el Control de Proyectos de TI en el
Instituto Nacional de Estadística e Informática, Lima ,2022**

TESIS PARA OBTENER EL GRADO ACADÉMICO DE:

Maestro en Ingeniería de Sistemas con Mención en Tecnologías de la Información

AUTOR:

Paredes Reyes Ivan Wilber (ORCID: 0000-0001-7162-0758)

ASESOR:

Dr. Acuña Benites, Marlon Frank (ORCID: 0000- 0001-5207-9353)

LÍNEA DE INVESTIGACIÓN:

Sistemas de Información y Comunicaciones

LÍNEA DE RESPONSABILIDAD SOCIAL UNIVERSITARIA:

Desarrollo económico, empleo y emprendimiento

LIMA – PERÚ

2022

Dedicatoria

A mi familia, mis padres, hermanas por su apoyo en varios aspectos, por su tiempo y esfuerzo que han dedicado en el logro de mi formación personal y profesional en cada momento.

Agradecimiento

Agradezco a mi familia que me alentaron y apoyaron en cada momento mi formación personal y profesional. Al Dr. Marlon Acuña Benites por compartir su experiencia profesional y su asesoría permanente para la culminación de la presente tesis.

Índice de contenidos

	Pg.
Carátula	i
Dedicatoria	ii
Agradecimiento	iii
Índice de contenidos	iv
Índice de tablas	v
Índice de figuras	vi
Resumen	vii
Abstract	viii
I. INTRODUCCIÓN	1
II. MARCO TEÓRICO	4
III. METODOLOGÍA	13
3.1. Tipo y diseño de investigación	13
3.2. Variables y operacionalización	13
3.3. Población, muestra y muestreo	14
3.4. Técnicas e instrumentos de recolección de datos	15
3.5. Procedimientos	16
3.6. Método de análisis de datos	16
3.7. Aspectos éticos	16
IV. RESULTADOS	17
V. DISCUSIÓN	28
VI. CONCLUSIONES	32
VII. RECOMENDACIONES	33
Referencias	34
ANEXOS	42

Índice de tablas

	Pg.
Tabla 1. Ficha técnica del instrumento	15
Tabla 2. Medidas descriptivas del indicador	17
Tabla 3. Pruebas de normalidad de los indicadores	19
Tabla 4. Índice de Eficacia interna de las actividades	20
Tabla 5. Porcentaje de actividades terminadas a tiempo	21
Tabla 6. Eficiencia de colaboradores	21
Tabla 7. Conclusiones de la prueba de normalidad	22
Tabla 8. Estadísticos descriptivos hipótesis general	23
Tabla 9. Pruebas de normalidad hipótesis general	23
Tabla 10. Medida descriptiva de la hipótesis general	24
Tabla 11. Medidas descriptivas del indicador	25
Tabla 12. Medidas descriptivas del indicador	26
Tabla 13. Prueba T Student	27

Índice de figuras

	Pg.
Figura 1. Medidas descriptivas del indicador	17
Figura 2. Medidas descriptivas del indicador	18
Figura 3. Medidas descriptivas del indicador	18

Resumen

En la gestión de proyectos sin adecuar el tiempo, recursos, costos ocasiona que se agregue carga laboral de lo planificado inicialmente. El propósito del trabajo de investigación fue determinar el uso de machine learning en el control de los proyectos en el Instituto Nacional de Estadística e Informática.

El diseño de investigación es experimental y presenta un enfoque cuantitativo. El tamaño de la muestra fue de 96 proyectos y se aplicó como instrumento de investigación la ficha de observación y para la prueba de hipótesis en el análisis inferencial se utilizó la prueba de Wilcoxon y la prueba T Student.

Se concluyó que la implementación de machine learning mejora significativamente el control de los proyectos, teniendo como mejora sus indicadores, demostrando que la eficacia interna de las actividades mejoro en un 52.59% y las actividades terminadas fuera de tiempo disminuyeron en 9.27% y la Eficiencia de colaboradores mejoro en un 32.92%. Después de implementar machine learning el control de proyectos mejoró de un 52.39% a un 60.01%. El valor de significancia fue 0.000, el cual es menor del valor del error 0.05, entonces, se rechaza la hipótesis nula (H_0) y se acepta la hipótesis alterna (H_a).

Palabras clave: Machine learning , Control de los proyectos , Indicadores ,Proyectos

Abstract

In the management of projects without adjusting the time, resources, costs, it causes the workload to be added to what was initially planned. The purpose of the research work was to determine the use of machine learning in the control of projects at the National Institute of Statistics and Informatics.

The research design is non-experimental and presents a quantitative approach. The sample size was 96 projects and the observation sheet was applied as a research instrument and the Wilcoxon test and the T Student test were used for the hypothesis test in the inferential analysis.

It was concluded that the implementation of machine learning significantly improves the control of the projects, having as an improvement its indicators, demonstrating that the internal efficiency of the activities improved by 52.59% and the activities completed out of time decreased by 9.27% and the Efficiency of collaborators improved by 32.92%. After implementing machine learning, project control improved from 52.39% to 60.01%. The significance value was 0.000, which is less than the error value 0.05, so the null hypothesis (H_0) is rejected and the alternate hypothesis (H_a) is accepted.

Keywords: Machine learning, Project control, Indicators, Projects

I. INTRODUCCIÓN

Para la administración de proyectos convergen varios elementos de toma de decisiones, psicología, dirección, gestión logística, gestión económica, conocimientos propios del área, tecnologías de información y matemáticas que permiten conseguir objetivos planteados en un tiempo determinado con recursos limitados sosteniendo un balance entre tiempo, costo y calidad. Aunque existen diversas escuelas de gestión de proyectos en el mundo, por ejemplo: ISO 21500, CMMI, PMBOK, Prince; la gran cantidad de proyectos renegociados y fallidos es muy alta. (Hasan y otros, 2020)

El alcance del proyecto sin adecuar costos, recursos, ajuste de tiempo ocasiona que no se materialice la utilidad del proyecto además agrega suma carga laboral de lo planificado inicialmente. De investigaciones pasadas se obtiene tres motivos principales de fracaso de los proyectos: Variación en las prioridades, variación de los objetivos y errores de compilación de requisitos. También la aparición de nuevas tecnologías y sistemas tienen un efecto en el alcance. (Project Management Institute, 2018)

En el Perú desde que se implementó el uso de la informática al entorno laboral ocasiona temores sobre la disminución del trabajo. En el Instituto Nacional de Estadística e Informática (INEI) se presentan dificultades para adaptarse al cambio por un grupo de personas que no se alejan de patrones establecidos. También el INEI es una institución técnica del gobierno peruano, con gestión y autonomía técnica que depende de la presidencia del Consejo de ministros y es responsable de coordinar, normar, dirigir y supervisar las tareas estadísticas del Perú. (Instituto Nacional de Estadística e Informática, 2017)

Con el avance actual de la tecnología, la información además de ser importante se ha incrementado considerablemente. Por la tanto para los tomadores de decisiones de las organizaciones, se requiere un acceso rápido en forma de reportes para poder identificar problemas y tomar acciones correctivas. La Oficina de Infraestructura (OI) del INEI tiene asignado en su plan estratégico

para los años 2016, 2017, 2018, 2019, 2020 aproximadamente 100 proyectos para lo cual necesita de una solución de machine learning para poder predecir el cumplimiento del proyecto.

El problema radica en que no existe un análisis histórico del volumen de datos almacenados lo cual ocasiona incertidumbre y demora para tomar decisiones al director técnico. No se cuenta con una herramienta informática que manipule los datos, los convierta en información útil, los procese y permita mostrarlos eficientemente. Además, no se utilizan indicadores para medir la gestión del área de infraestructura.

La importancia es contar con información automatizada para el control de los proyectos, tomar decisiones para resolver los problemas y reasignar recursos y realizar seguimiento presupuestal. La finalidad de esta aplicación es para que el director técnico de la OI pueda hacer seguimiento integrado de los proyectos y también tener ratios de cumplimiento de los proyectos.

En tal sentido, el presente trabajo se argumentará la pregunta: ¿De qué forma el machine learning mejora el control de proyectos de TI en el INEI, Lima ,2022?

También los problemas específicos, se mencionan a continuación: (a) ¿De qué forma machine learning mejora la eficacia de los proyectos de TI en el INEI, Lima ,2022?; (b) ¿De qué forma machine learning mejora la productividad de los proyectos en el INEI,2022?;(c) ¿De qué forma machine learning mejora el desempeño de los proyectos de TI en el INEI, Lima ,2022?

Por otro lado, para alcanzar los objetivos que se encuentran planteados en el problema se estableció el siguiente objetivo general: Establecer que machine learning mejora el control de proyectos de TI en el Instituto Nacional de Estadística e Informática. Lima ,2022. Asimismo, los objetivos específicos son los siguientes: (a) Establecer que machine learning aumenta la eficacia de los proyectos de TI en el INEI. Lima ,2022, (b) Establecer que machine learning aumenta la productividad de los proyectos en el INEI,Lima ,2022, (c) Establecer que machine learning aumenta el desempeño de los proyectos de TI en el INEI,Lima ,2022.

El presente trabajo de investigación tiene tres justificaciones los cuales mencionamos a continuación: Justificación teórica, se fundamenta en la masificación del uso de tecnologías de inteligencia artificial como machine learning para mejorar la gestión de proyectos. Justificación practica su aplicación mejoraría los tiempos de ejecución de los proyectos alertando la situación actual de los proyectos a los colaboradores de la oficina de infraestructura del INEI. Justificación metodológica tendrá resultados mediante el uso de la inteligencia artificial que encontrará vínculos entre las variables de los proyectos y de machine learning.

De igual modo se planteó la siguiente hipótesis general: Machine learning aumenta significativamente el control de proyectos de TI en el INEI, Lima ,2022. Respecto a las hipótesis específicas, tenemos las siguientes: (a) Machine learning aumenta la eficacia de los proyectos de TI en el INEI, Lima ,2022, (b) Machine learning aumenta la productividad de los proyectos en el INEI ,Lima ,2022, (c) Machine learning aumenta el desempeño de los proyectos de TI en el INEI, Lima ,2022.

II. MARCO TEÓRICO

En los antecedentes nacionales para sustentar el trabajo de investigación tenemos:

Según (Álvarez Ochoa, 2021) determinó que existe relación entre los proyectos de inversión pública en el Perú e inteligencia artificial, pues demostró que el algoritmo de optimización contiene parámetros de metodología que lo acreditan. Pues se reconocieron trece factores críticos para la gestión de proyectos de saneamiento que mediante un análisis de inteligencia artificial mejora la efectividad de la toma de decisiones.

Para (León, 2020) en su investigación sobre machine learning y Business Intelligence pretende dar a conocer un modelo de solución del consumo de agua para el control de calidad y monitoreo en el Distrito de Breña con el fin de obtener mejores decisiones , analizando factores asociados en los suministros en el Centro de lecturas de medidores de Breña en el cual se monitorea en tiempo real con un análisis predictivo para alertar de al personal y tomen las medidas correctivas sobre el consumo de agua no facturada para minimizar los reclamos sobre facturación y mejorar el servicio al cliente.

Según (Silva, 2021) en su trabajo planteó como propósito determinar cómo machine learning permite mejorar las operaciones comerciales en la empresa Redondos. La metodología que utilizó fue la recopilación de datos mediante fichas de observación. La principal conclusión es que machine learning mejora los indicadores de las operaciones comerciales.

Así también (López J. , 2019) en su trabajo de investigación cuantitativa y como instrumento la observación y ficha de recolección sobre instituciones financieras menciona que existe mayor riesgo de incumplimiento de pagos en una microfinanza que en una banca tradicional. Para reducir el riesgo de clientes morosos, se le solicita información económica en detalle. El propósito principal del trabajo fue comparar diferentes algoritmos para reportar con exactitud la predicción de incumplimiento de pagos en microfinanzas.

(Coronel, 2021) sostiene la necesidad de mejorar la selección de personal administrativo, pues se pierde dinero y tiempo, ya que continuamente se debe

realizar nuevas convocatorias para cubrir los puestos laborales. Entonces el autor utilizó algoritmos de machine learning para optimizar la selección de empleados administrativo.

(Yamao, 2018) en su trabajo del tipo cuantitativa sobre la predicción de las notas de alumnos que recién ingresan a la facultad de ingeniería de una universidad particular. El autor utilizó variables como sociales, académicos y económicos. La conclusión más importante del trabajo de investigación es que mediante modelos predictivos se logró identificar que alumnos presentarían problemas al iniciar estudios universitarios.

(Quipas, 2021) en su trabajo tuvo como propósito precisar si la aplicación de machine learning mejora la gestión del conocimiento en Concytec. En el trabajo se utilizó indicadores de tiempo para comparar la gestión del conocimiento tradicional y con el uso de machine learning.

En cuanto a los antecedentes internacionales, el investigador (Caro y otros, 2021) en su estudio sobre un modelo de evaluación de proyectos menciona que es necesario aplicar nuevas tecnologías en la fase de planeación para predecir el éxito o fracaso de un proyecto. El autor concluye que la implementación de machine learning proporciona para la evaluación de proyectos un indicador que define las métricas financieras utilizadas comúnmente para tomar decisiones, esto ayuda a reducir la cantidad de proyectos fracasados.

(Maher, 2021) en su investigación sobre aplicación de la inteligencia artificial en gestión de proyectos, dice que los gerentes mediante técnicas de análisis predictivo tienen nuevos métodos y técnicas para que sus labores sean más eficientes y rápidas. El trabajo tuvo como objetivo analizar la inteligencia artificial en la gestión de proyectos en tiempo real y la manera de tomar decisiones por parte del gerente del proyecto.

(Najah y otros, 2021) indica que para mejorar la eficiencia y mantener la sostenibilidad de un proyecto de software son obstáculos a los que se enfrentan los gestores de proyectos. La probabilidad de fracaso del proyecto es generalmente debido a la falta de conocimientos, habilidades, recursos y

tecnología durante el desarrollo del proyecto. Para implementar su investigación se obtuvieron el conocimiento de datos históricos para el modelo predictivo utilizando matemáticas incluyendo regresión lineal y estudio de asociación o enfoques de aprendizaje automático (ML), como la Red de Redes Artificiales (ANN) y las Máquinas de Vectores de Soporte (SVM) proporcionan un método que se centra en la evidencia actual e histórica del proyecto para pronosticar el futuro del proyecto.

(Pan & Stark, 2022) desarrolló una investigación cuyo fin principal es estudiar el estado actual para la predicción de la mantenibilidad del software y descubrir el progreso realizado, las limitaciones y los desafíos, junto con los requisitos de investigación futuros. Este estudio clasifica el concepto de mantenibilidad del software según tres dimensiones: la medida de la mantenibilidad (variable dependiente), las métricas consideradas (variables independientes) y los modelos empleados. Según (Banco Interamericano de Desarrollo, 2018) la predicción es una parte central de la estimación, que es un aspecto crucial de la planificación del proyecto e implica la determinación de varios factores que implican la duración, el personal, el tamaño, los costos y el esfuerzo. La predicción depende principalmente de los atributos históricos de calidad interna y externa de los proyectos terminados.

Para (Lucero, 2022) la correlación entre los atributos internos, "variables independientes", y los atributos externos, "variables dependientes", es una parte reconocida de la predicción de la mantenibilidad del software. Los modelos de predicción de mantenibilidad de software se han investigado para ayudar a las organizaciones a utilizar los costos, asignar recursos y obtener un plan de gestión preciso y un proceso de mantenimiento eficaz. Sin embargo, la predicción de la mantenibilidad del software es una tarea desafiante y requiere modelos de predicción precisos.

(Musbah y otros, 2022) en su trabajo de investigación indica que los sistemas de energía híbridos son una de las soluciones efectivas para la demanda de energía, especialmente en zonas remotas. Incluyen procedencia de energía renovables como la eólica y/o solar. Las energías renovables son intermitentes, fluctuantes y no lineales. Entonces, una gestión energética eficaz

juega un rol crucial para la organización del flujo de energía en las fuentes de energía híbridas. La gestión de la energía híbridos y los recursos energéticos tradicionales se realiza mediante el uso de algún software como MATLAB, Simulink, o inteligencia artificial como el algoritmo de evolución diferencial, el modelo de serie temporal de estado casi estacionario de lógica difusa, la optimización de enjambre de partículas, la optimización genética. algoritmo y redes neuronales artificiales. En el presente trabajo, el plan de gestión energética se pronostica utilizando un conjunto de datos.

(Garbero y otros, 2021) menciona en su investigación que los avances en el aprendizaje automático y estudios de Big Data tienen un gran potencial para que los organismos internacionales de desarrollo aprovechen la amplia información generada a partir de rendición de cuentas para obtener nuevos conocimientos, proporcionando análisis que mejoran la toma de decisiones. Desde un análisis de gestión del conocimiento, los informes operativos del proyecto son registros históricos importantes, ya que cuentan la historia de un proyecto, captan datos generados a lo largo del ciclo del proyecto, analizan los logros de los objetivos de desarrollo previstos y generan aprendizaje que puede informar para operaciones futuras. Para (Gonzalez y otros, 2019) la minería de texto descriptiva es un enfoque que pretende automatizar la obtención de información de alta calidad del texto, encontrando patrones y tendencias a través del aprendizaje automático, las estadísticas y la lingüística. Esta técnica permite analizar grandes cantidades de texto no estructurado para obtener patrones implícitos, previamente desconocidos y muy útiles para convertirse en inteligencia importante para la toma de decisiones.

Para (Nwanyanwu, 2018) indica que el control de fraudes consiste en utilizar sistemas de control interno y leyes, a través del establecimiento de auditoría interna, auditoría externa, oficinas antifraudes y comités de auditoría. El uso de controles disminuye el fraude y mejora la confianza de los inversionistas de una organización. El autor menciona que las leyes son vitales, pues sancionar a los infractores es una forma de disuadir a otros a cometer delitos. También en una organización se deben implantar castigos como despido

del trabajo, suspensión del empleado o el pago de las cantidades cometidas en el delito.

(A.S. & Edayadiyil, 2022) menciona en su investigación que el seguimiento y control del progreso del proyecto es crítica para la gestión de proyectos de construcción. Las prácticas actuales para pronosticar la efectividad de un proyecto de construcción incluyen una evaluación laboriosa y propensa a errores. La digitalización puede mejorar las prácticas de gestión para el campo de la construcción que es reacio a admitir nuevas tecnologías. Esto señala la importancia de una manera automatizada para calcular el progreso de los proyectos. El propósito de la investigación es implementar un método de monitoreo para el Seguimiento del Progreso de la Construcción. Se utilizan técnicas de aprendizaje automático y procesamiento de imágenes para garantizar Seguimiento del Progreso de la Construcción automáticamente. Se centra la actividad de albañilería en la edificación, también puede ampliarse a otros trabajos de construcción. El modelo de aprendizaje profundo que se propone se entrena con un conjunto de datos de 356 imágenes tomadas de varios lugares de trabajo y páginas web. El sistema propuesto utiliza el lenguaje Python y genera un informe de progreso automatizado en Microsoft Excel.

(Mis & Sankaranarayanan, 2021) en su trabajo de investigación analiza las métricas e indicadores utilizados junto con modelos de software para identificar el porcentaje de riesgo que conduce a lograr el éxito en el proceso de desarrollar software. Se obtiene mediante la elaboración de una encuesta de datos recibida por parte de expertos en desarrollo de software. Entonces, se estableció un sistema de predicción donde se entrenaron algoritmos de aprendizaje automático, tales como máquina de vectores de soporte, k-vecino más cercano, red neuronal artificial y bosque aleatorio para predecir el modelo de software con sus respectivos porcentajes de riesgo. Las principales contribuciones del trabajo de investigación son las siguientes, Encuesta con cuestionario que tiene métricas e indicadores para identificar el modelo de software análisis comparativo de diferentes modelos de machine learning para clasificación de riesgos. Cálculo del porcentaje de riesgo basado en la clasificación de riesgo para varios modelos de software.

Entre las teorías generales para la elaboración del trabajo de investigación tenemos a la inteligencia artificial y la administración.

Según (Valbuena, 2021) la inteligencia artificial estudia probabilidad elaborar maquinas con pensamiento que se apoyan en modelos bioma temáticos que pueden desempeñar similares labores que los seres humanos. Pues la inteligencia artificial pertenece a las disciplinas de las ciencias computacionales y las neurociencias cognitivas, así como a la detección de patrones y modelos de razonamiento con incertidumbre.

La inteligencia artificial es la ciencia que fabrica maquinas inteligentes mediante programas que poseen inteligencia informática; además se relaciona con la tarea de utilizar computadoras para entender la inteligencia humana. Actualmente la inteligencia artificial puede aumentar o reemplazar la experiencia requerida para la toma de decisiones. (Srikrishna & Shivendra, 2022)

La administración es un conjunto de herramientas, técnicas y conceptos sobre la forma de ejecutar los proyectos dentro del presupuesto, tiempo y según lo solicitado por el cliente en el contexto de la estrategia empresarial. (Vrchota & Rehor, 2021)

Para (Mero, 2018) la administración tiene como función básica dirigir la empresa a un logro mediante la selección de objetivos apropiados. Henri Fayol detallo actividades para controlar, planear, organizar, dirigir, coordinar los procesos administrativos. Varios investigadores reconocen a la administración como una ciencia social que posee técnicas, principios y practicas a grupos humanos que tienen metas comunes que personalmente no es posible lograr.

Las teorías que sirven de base para el presente trabajo de investigación son:

(Kumar y otros, 2018) indica que machine learning es una especialidad de las ciencias de la computación muy importante, pues concede a las computadoras la capacidad de aprender sin que se les enseñe explícitamente. Por lo tanto, los algoritmos de aprendizaje automático están creados en base al proceso de aprendizaje humano, ejercitando de forma iterativa a partir de los

datos proporcionados, permitiendo que las computadoras localicen información oculta mediante el reconocimiento de patrones.

(Chio & Freeman, 2018) menciona que el aprendizaje automático se basa en procesos y algoritmos que aprenden de experiencias y datos anteriores para predecir efectos futuros. Entonces machine learning es un conjunto de métodos que están desarrolladas en sistemas informáticos, que mediante la minería de datos descubren patrones y deducciones con el uso de los datos proporcionados.

(Hurwitz & Kirsch, 2018) indican que machine learning aprende cuando se ingresan los datos y no se ejercita de la programación habitual, también es útil descubrir modelos con mayor precisión con el uso de datos. La solución obtenida permite comprender un prototipo de machine learning usando el correspondiente algoritmo que se alimenta con el ingreso de datos capturados. Después se recibe un bosquejo mediante la incorporación de datos, obteniendo como respuesta un algoritmo de machine learning que nos posibilita predecir información futura.

(Krishna y otros, 2021) sostiene que machine learning forma parte de la inteligencia artificial, cuya meta es desarrollar programas sin programación directa, pues se basa en el conocimiento para el desarrollo de programas. En conclusión, machine learning permite al computador prepararse y modificar su conducta sin la supervisión de una persona.

Según (Kashyap, 2017) mediante un estudio vinculó las tecnologías informáticas con machine learning para tomar decisiones que se realiza en el momento y tiempo real necesitan de algoritmos de aprendizaje automático cuya precisión es de 99%.

Para (protiviti, 2019) los gerentes que no manejan la tecnología se muestran reacios a la aplicación de machine learning, pues es menos riguroso al implementarse generando desconfianza.

Para (Mora, 2018) machine learning usa gran cantidad de datos para nutrir a los algoritmos y para su aprendizaje para predecir o recomendar sobre un tema. También se destaca en descubrir conocimientos e identificar patrones que no se encuentran visibles en un conjunto de datos.

(Alzubi y otros, 2018) menciona en su trabajo que los términos de Inteligencia Artificial y machine learning no son nuevas. Han sido estudiados, aplicados y mejorados por ingenieros, investigadores, estudiantes y profesionales durante más de 60 años. La base matemática del aprendizaje automático radica en la estadística, la probabilidad y el álgebra. El desarrollo del machine learning y la inteligencia artificial empezó en las décadas de 1950 y 1960 con las investigaciones de Alan Turing, Frank Rosenblatt, Arthur Samuels entre otros. Samuels, propuso el primer modelo de aprendizaje automático logrando que aprendiera un ordenador IBM 701 un juego de damas. Rosenblatt creó el Perceptrón, un algoritmo muy conocido de aprendizaje automático basado en neuronas biológicas que estableció las bases de la red neuronal artificial.

(Sazili y otros, 2019) en su investigación expone que factores influyen en el éxito de proyectos de tecnología de información como la gestión de: calidad, personas, alcance, costos, tiempo. El éxito de la entrega del proyecto se realiza a través de una encuesta en varios departamentos de la industria financiera de Malasia. El cuestionario de la encuesta fue utilizado como técnica para la recolección de datos. Para (Andrade y otros, 2019) en la estadística y la observación evidenciaron que el manejo adecuado de alcance, tiempo; el costo, la calidad y el personal tienen una sólida influencia en el éxito de los proyectos.

(Sorooshian & Yin, 2020) menciona en su trabajo de investigación que los proyectos de tecnologías de información tienen innumerables riesgos que son complicados de administrar y mitigar para el gerente del proyecto, que a menudo se enfrenta la tarea de pronosticar sobre los riesgos potenciales que puede amenazar la implementación de un proyecto de tecnologías de información. Este estudio tuvo como objetivo estudiar los factores de riesgo críticos que enfrentan los proyectos de TI para ayudar al gerente del proyecto en la disminución de los problemas de gestión actual. Para (Rudas, 2017) el éxito de un proyecto depende del total uso del sistema, también del correcto manejo de la gestión de riesgos, donde el gerente de TI busca gestionar el equilibrio entre el coste y el retorno de la inversión. Para (Vigo y otros, 2018) la identificación del riesgo es un factor importante para que las empresas reduzcan la probabilidad del impacto

de amenazas en los proyectos de TI, así como en la previsión de posibilidades de riesgo.

Dentro de las dimensiones del trabajo de investigación tenemos a la eficacia que radica en la disposición de la organización para lograr los propósitos y lograr la satisfacción del cliente estableciendo correctamente sus expectativas y necesidades. Esto se logra deduciendo las características que tendrán los productos para que cumplan los requisitos del cliente. (García y otros, 2019)

Otra dimensión que se consideró para el trabajo de investigación es el desempeño laboral que se relaciona con la habilidad que tiene la persona para generar valor a la empresa en situaciones dinámicas. El empleado debe conseguir cumplir con la misión, metas y actividades para el equilibrio de las necesidades de los interesados. (López y otros, 2022)

Finalmente, la última dimensión considerada en el trabajo de investigación es la eficiencia que se emplea para medir el cumplimiento de actividades y el uso de recursos mediante un nexo entre cantidad de recursos utilizados y cantidad de recursos programados. Además, la eficiencia también se relaciona con la calidad que se obtiene al aprovechar los recursos utilizados al transformarse en servicios y productos. (Duque & Supo, 2021)

III. METODOLOGÍA

3.1. Tipo y diseño de investigación

Tipo de investigación

El trabajo de investigación es aplicada, pues se trata de determinar la mejora del control de proyectos mediante machine learning, pues la manera tradicional es mediante la experiencia humana que puede presentar limitaciones cognitivas, pero en machine learning no presenta sesgos cognitivos. También la tesis es cuantitativa, pues su propósito es de identificar patrones, evaluar variables, medir resultados y validar teorías. (Hernández y otros, 2014)

Diseño de investigación

La investigación tiene diseño experimental pues se manipula la variable independiente para estudiar los efectos en la variable dependiente. También es del tipo preexperimental, pues el control de las variables es mínimo. En el presente trabajo se utilizó un pretest, postest y con un total de 96 muestras. (Hernández y otros, 2014)

3.2. Variables y operacionalización

Definición conceptual de la variable independiente: Machine learning

Machine learning es un grupo de técnicas que utiliza algoritmos de inteligencia artificial para extraer patrones mediante métodos estadísticos, matemáticos para optimizar y descubrir conocimientos, también se divide en tres aprendizajes: supervisado, no supervisado y por refuerzo. (Akbar y otros, 2021)

Definición conceptual de la variable dependiente: Control de proyectos

El control de proyectos es una de las funciones críticas en la gestión de proyectos, por lo tanto, implica varias decisiones como problemas para programar los proyectos, informar del estado del proyecto comparándolo con el plan inicial, detección de situaciones fuera de control, tomar decisiones apropiadas y analizar situaciones. También según el autor indica que la mayoría de los problemas en el control de proyectos se relacionan con la programación de proyectos y asignación de recursos. (Pellerin & Perrier, 2019)

Definición operacional de la variable dependiente: Control de proyectos

Para el bosquejo de Control de proyectos se ha estimado tres indicadores de medición: (a) Eficacia interna de las actividades, (b) Actividades terminadas fuera de tiempo (c) Porcentaje de proyectos terminados a tiempo. Siendo la unidad de medida para los tres indicadores el porcentaje. En los tres indicadores se empleó como instrumento de recolección de datos a la guía de observación.

3.3. Población, muestra y muestreo

Población

Es un conjunto de todas las observaciones que tienen en común definidas especificaciones. También se tienen que especificar por sus características de contenido, tiempo y lugar y su delimitación depende de los propósitos planteados en el estudio. En el presente trabajo de investigación la población está conformada por 284 observaciones para calcular los datos del pretest y postest. (Hernández y otros, 2014)

Muestra

La muestra es una porción de la población que tiene características comunes. En la realidad muchas veces es imposible medir a toda la población por lo que se selecciona a un subconjunto de elementos representativos que es la muestra. (Otzen & Manterola, 2017)

Procedemos a estimar la proporción de la muestra:

Empleando la formula:

$$n = \frac{284 * 1.96^2 * 0.05 * 0.95}{0.05^2 * (284 - 1) + 1.96^2 * 0.05 * 0.95} = 96$$

Se obtiene un muestrario de 96 observaciones a partir de una población de 284 proyectos.

Muestreo

El muestreo que utilizaremos en el estudio es el muestreo probabilístico aleatorio simple que permite que individuos de una población objetivo tengan igual oportunidad de pertenecer a la muestra. Quiere decir que la posibilidad de ser escogido para la muestra un sujeto x es autónomo del resto de sujetos que pertenecen a la población blanco. (Otzen & Manterola, 2017)

3.4. Técnicas e instrumentos de recolección de datos

Técnicas de recolección de datos

La observación se debe realizar de manera sistemática y con una correcta metodología planificada, en un entorno que se pueda controlar para relacionarlas con las teorías del trabajo de investigación. En nuestro caso utilizaremos a la observación de técnica de recolección de datos para que nos facilite la recolección y el proceso de la información. (Hernández & Duana, 2020)

Instrumentos de recolección de datos

En el presente estudio se empleó como instrumento de recolección de datos la Guía de observación de campo, para medir el pretest y postest. Según (Hernández y otros, 2014) la herramienta de medición de recolección de datos son técnicas que permiten el registro de datos de la realidad observable.

Tabla 1

Ficha técnica del instrumento

Nombre del instrumento	Guía de observaciones de medición del indicador
Autor:	Recolección y procesamiento
Año:	Movilidad
Descripción:	Fuentes bibliográficas
Tipo de instrumento:	Guía de observación de campo
Objetivo:	

Indicadores:	Determinar la mejora del machine learning en el control de proyectos de TI en el Instituto Nacional de Estadística e Informática. Lima ,2022 a) Eficacia interna de las actividades b) Eficiencia global de equipos c) Porcentaje de proyectos terminados a tiempo
Número de observaciones a recolectar:	
Aplicación:	96
	Directa

Fuente: Elaboración propia

3.5. Procedimientos

Se realizó los siguientes pasos en el trabajo de investigación:

Se recolecto información en el INEI ubicado en la ciudad de Lima, con una muestra de 96 proyectos, para obtener los datos se utilizó la técnica de observación como instrumento a la ficha de registro.

3.6. Método de análisis de datos

Para el estudio de los datos del trabajo de investigación del pretest y el postest se empleó Microsoft Excel y software estadístico SPSS v25. También en el análisis descriptivo se usó figuras y tablas, describiendo medidas de tendencia central empleando la media para interpretar los datos del instrumento.

3.7. Aspectos éticos

Con el objetivo de cumplir con la integridad del trabajo de investigación se cumplió con las normas establecidas por la Universidad Cesar Vallejo mediante la Resolución del Vicerrectorado de investigación N°110-2022-VI-UCV que contiene las directivas para elaborar la presente investigación. También para respetar las políticas anti-plagio se empleó el software Turnitin en varias oportunidades. Además, para presentar un trabajo ordenado para consultar los documentos empleados se utilizó las normas APA 7.

IV. RESULTADOS

Análisis descriptivo

En la tabla 2, se visualiza la estadística descriptiva de la Eficacia interna de las actividades (I1), Actividades terminadas fuera de tiempo (I2) y Eficiencia de colaboradores (I3) antes y después de implementar el machine learning.

Tabla 2

Medidas descriptivas del indicador

Estadísticos descriptivos					
	N	Mínimo	Máximo	Media	Desv. Desviación
I1Pretest	96	,17	,53	,3534	,09220
I1Posttest	96	,71	,99	,8793	,06904
I2Pretest	96	,47	,96	,7159	,12933
I2Posttest	96	,01	,29	,0927	,07615
I3Pretest	96	,27	,67	,5007	,08584
I3Posttest	96	,64	,99	,8299	,07786
N válido (por lista)	96				

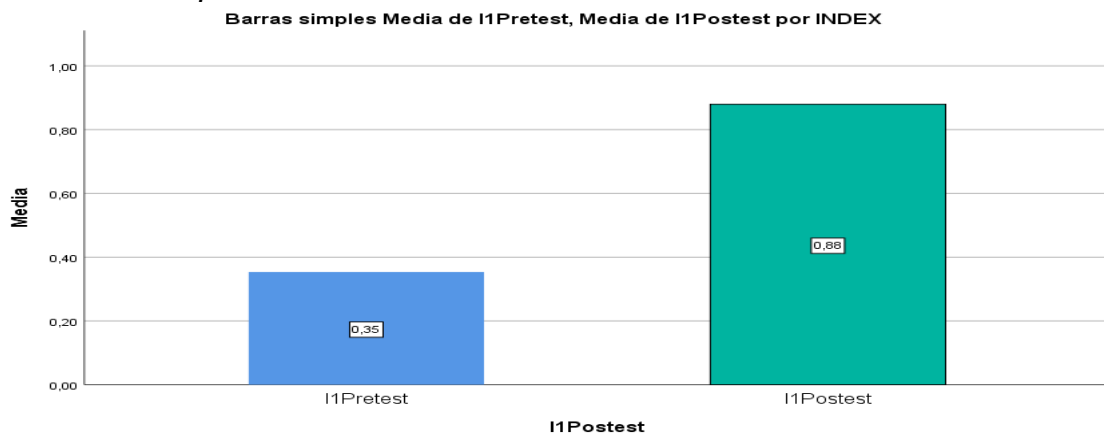
Fuente: Elaboración propia

Indicador 1: Eficacia interna de las actividades

En la figura 1, se visualiza el comportamiento del indicador Eficacia interna de las actividades, antes y después de la implementación del machine learning en base a los datos obtenidos en la ficha de observación, por lo cual, se puede concluir que el índice de Eficacia interna de las actividades mejoro de un 35.34% a un 87.93%. En este indicador la mejora se dio en el aumento del 52.59%.

Figura 1

Medidas descriptivas del indicador

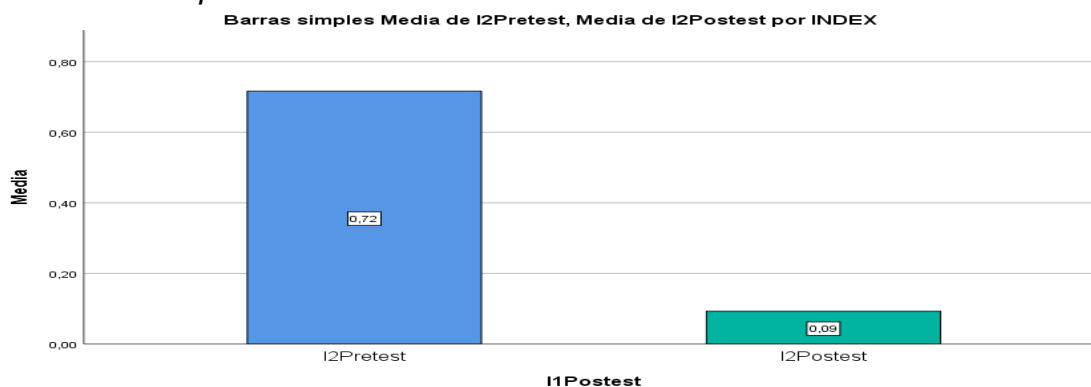


Fuente: Elaboración propia

Indicador 2: Actividades terminadas fuera de tiempo

En la figura 2, se visualiza el comportamiento del indicador actividades terminadas fuera de tiempo, antes y después de la implementación del machine learning en base a los datos obtenidos en la ficha de observación, por lo cual, se puede concluir que el índice de actividades terminadas fuera de tiempo mejoro de un 71.59% a un 9.27%. En este indicador la mejora se dio en la disminución del 62.32%.

Figura 2
Medidas descriptivas del indicador

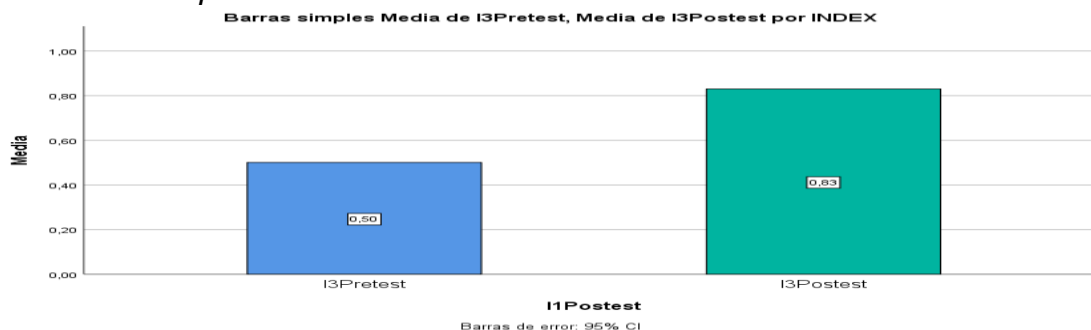


Fuente: Elaboración propia

Indicador 3: Eficiencia de colaboradores

En la figura 3, se visualiza el comportamiento del indicador Eficiencia de colaboradores, antes y después de la implementación del machine learning en base a los datos obtenidos en la ficha de observación, por lo cual, se puede concluir que el índice de Eficiencia de colaboradores mejoro de un 50.07% a un 82.99%. En este indicador la mejora se dio en el aumento del 32.92%.

Figura 3
Medidas descriptivas del indicador



Fuente: Elaboración propia

Análisis inferencial

Prueba de normalidad

Se aplicó con el procedimiento de Kolmogorov-Smirnov, pues la muestra es superior a 30; para este método se utilizó el software IBM SPSS v25, con un nivel de confiabilidad del 95%, entonces si el valor de significancia es menor a 0,05 es una distribución no normal.

A continuación, en la tabla 3 se observan los resultados de las pruebas de normalidad antes y después de implementar machine learning.: Eficacia interna de las actividades, Porcentaje de actividades terminadas a tiempo y Eficiencia de colaboradores

Tabla 3

Pruebas de normalidad de los indicadores

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Estadístico	gl	Sig.	Estadístico	gl	Sig.
I1Pretest	,122	96	,001	,961	96	,006
I1Postest	,133	96	,000	,950	96	,001
I2Pretest	,154	96	,000	,941	96	,000
I2Postest	,211	96	,000	,872	96	,000
I3Pretest	,071	96	,200*	,982	96	,197
I3Postest	,076	96	,200*	,978	96	,103

*. Esto es un límite inferior de la significación verdadera.

a. Corrección de significación de Lilliefors

Fuente: Elaboración propia

En tabla 3 se evidenció que el índice de Eficacia interna de las actividades para el pretest no cumplió con la normalidad, debido a que su p valor (0.001) no superó el 5% de margen de error. También, el postest su p valor (0.000) tampoco superó el 5% de margen de error, lo que evidencia la no normalidad. Al respecto del índice Porcentaje de actividades terminadas a tiempo, en el pretest su p valor (0.000) no superó el 5% de margen de error que determina que no existe normalidad, también, en el postest el p valor (0.000) no superó el 5% de margen de error, lo que indicó la existencia de no normalidad. En relación con el índice de Eficiencia de colaboradores en el pretest su p valor (0.200) superó el 5% de

margen de error que conlleva al cumplimiento de la normalidad, del mismo modo en el posttest su p valor (0.200) supero el 5% de margen de error, por tanto, hay normalidad.

Prueba de normalidad: Indicador 1

En los resultados de la tabla 4 se evidenció que el índice de Eficacia interna de las actividades en el pretest no cumplió con la normalidad, debido a que su p valor (0.001) no superó el 5% de margen de error. Asimismo, en el posttest su p valor (0.000) tampoco superó el 5% de margen de error, lo que evidencia la no normalidad.

Tabla 4

índice de Eficacia interna de las actividades

	Pruebas de normalidad					
	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Estadístico	gl	Sig.	Estadístico	gl	Sig.
I1Pretest	,122	96	,001	,961	96	,006
I1Postest	,133	96	,000	,950	96	,001

*. Esto es un límite inferior de la significación verdadera.

a. Corrección de significación de Lilliefors

Fuente: Elaboración propia

Prueba de normalidad: Indicador 2

Al respecto del índice Porcentaje de actividades terminadas a tiempo en la tabla 5, se evidencia que el pretest su p valor (0.000) no superó el 5% de margen de error que determina que no existe normalidad, también, en el posttest el p valor (0.000) no superó el 5% de margen de error, lo que indicó la existencia de no normalidad.

Tabla 5
Porcentaje de actividades terminadas a tiempo

Pruebas de normalidad

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Estadístico	gl	Sig.	Estadístico	gl	Sig.
I2Pretest	,154	96	,000	,941	96	,000
I2Postest	,211	96	,000	,872	96	,000

*. Esto es un límite inferior de la significación verdadera.

a. Corrección de significación de Lilliefors

Fuente: Elaboración propia

Prueba de normalidad: Indicador 3

En relación con el índice de Eficiencia de colaboradores en la tabla 6 se evidencia que en el pretest su p valor (0.200) superó el 5% de margen de error que conlleva al cumplimiento de la normalidad, del mismo modo en el postest su p valor (0.200) superó el 5% de margen de error, por tanto, hay normalidad.

Tabla 6
Eficiencia de colaboradores

Pruebas de normalidad

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Estadístico	gl	Sig.	Estadístico	gl	Sig.
I3Pretest	,071	96	,200*	,982	96	,197
I3Postest	,076	96	,200*	,978	96	,103

*. Esto es un límite inferior de la significación verdadera.

a. Corrección de significación de Lilliefors

Fuente: Elaboración propia

En la siguiente tabla 7 se resume la técnica estadística a utilizar según los resultados de la prueba de normalidad.

Tabla 7*Conclusiones de la prueba de normalidad*

Indicador	Técnica estadística
Eficacia interna de las actividades	Wilcoxon
Porcentaje de actividades terminadas fuera de tiempo	Wilcoxon
Eficiencia de colaboradores	Prueba T

Fuente: Elaboración propia

Pruebas de hipótesis

Después de analizar la variable independiente, variable dependiente con sus dimensiones, se realizará la prueba de hipótesis donde identificaremos los valores que se obtienen, también aprobaremos la hipótesis inicial o la hipótesis nula.

Prueba de Hipótesis general: Mejora significativamente el control de proyectos de TI

En la tabla 8, se visualiza el comportamiento del promedio del pretest y postest, antes y después de la implementación del machine learning en base a los datos obtenidos en la ficha de observación, por lo cual, se puede concluir que mejora significativamente el control de proyectos de TI de un 52.39% a un 60.01%.

Tabla 8*Estadísticos descriptivos hipótesis general*

Estadísticos descriptivos					
	N	Mínimo	Máximo	Media	Desv. Desviación
pretest_general	96	,42	,67	,5239	,05155
posttest_general	96	,49	,66	,6001	,03246
N válido (por lista)	96				

```
EXAMINE VARIABLES=pretest_general posttest_general
```

Fuente: Elaboración propia

En los resultados de la tabla 9 se evidenció que la hipótesis control de proyectos no cumplió con la normalidad, debido a que su p valor (0.001) no superó el 5% de margen de error. Pero, en el posttest su p valor (0.22) superó el 5% de margen de error, lo que evidencia la no normalidad de la hipótesis general.

Tabla 9
Pruebas de normalidad hipótesis general

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Estadístico	gl	Sig.	Estadístico	gl	Sig.
pretest_general	,126	96	,001	,964	96	,010
posttest_general	,099	96	,022	,969	96	,022

a. Corrección de significación de Lilliefors

Fuente: Elaboración propia

Formulación de hipótesis estadística:

Ho: Machine learning no mejora significativamente el control de proyectos de TI en el Instituto Nacional de Estadística e Informática. Lima ,2022

Ha: Machine learning mejora significativamente el control de proyectos de TI en el Instituto Nacional de Estadística e Informática. Lima ,2022

El resultado de la prueba de normalidad de la hipótesis general no se distribuye normalmente, entonces se utilizó la prueba de Wilcoxon. También en la tabla 10 se muestra el valor Z que fue de -7.929, el cual se ubica en la zona de rechazo de la hipótesis nula (Ho) y el valor de significancia fue 0.000, el cual es menor del valor del error 0.05, por lo tanto, se rechaza la hipótesis nula (Ho) y se acepta la hipótesis alterna (Ha).

Tabla 10*Medida descriptiva de la hipótesis general***Prueba de rangos con signo de Wilcoxon**

		Rangos		
		N	Rango promedio	Suma de rangos
postest_general - pretest_general	Rangos negativos	9 ^a	17,78	160,00
	Rangos positivos	87 ^b	51,68	4496,00
	Empates	0 ^c		
	Total	96		

a. postest_general < pretest_general

b. postest_general > pretest_general

c. postest_general = pretest_general

Estadísticos de prueba^a

	postest_gene ral - pretest_gene ral
Z	-7,929 ^b
Sig. asintótica(bilateral)	,000

a. Prueba de rangos con signo de Wilcoxon

b. Se basa en rangos negativos.

Fuente: Elaboración propia

Prueba de Hipótesis específica 1: Eficacia interna de las actividades

Formulación de hipótesis estadística:

Ho: El machine learning no mejora significativamente la Eficacia interna de las actividades en el Instituto Nacional de Estadística e informática, Lima -2022.

Ha: El machine learning mejora significativamente la Eficacia interna de las actividades en el Instituto Nacional de Estadística e informática, Lima - 2022.

El resultado de la prueba de normalidad del indicador eficacia interna de las actividades no es normal se utilizó la prueba de Wilcoxon. También en la tabla 11 se muestra el valor Z que fue de -8.508, el cual se ubica en la zona de rechazo de la hipótesis nula (Ho) y el valor de significancia fue 0.000, el cual es menor del valor del error 0.05, por lo tanto, se rechaza la hipótesis nula (Ho) y se acepta la hipótesis alterna (Ha).

Tabla 11

Medidas descriptivas del indicador

Prueba de rangos con signo de Wilcoxon

		Rangos		
		N	Rango promedio	Suma de rangos
I1Postest - I1Pretest	Rangos negativos	0 ^a	,00	,00
	Rangos positivos	96 ^b	48,50	4656,00
	Empates	0 ^c		
	Total	96		

a. I1Postest < I1Pretest

b. I1Postest > I1Pretest

c. I1Postest = I1Pretest

Estadísticos de prueba^a

	I1Postest - I1Pretest
Z	-8,508 ^b
Sig. asintótica(bilateral)	,000

a. Prueba de rangos con signo de Wilcoxon

b. Se basa en rangos negativos.

Fuente: Elaboración propia

Prueba de Hipótesis específica 2: Porcentaje de Actividades terminadas fuera de tiempo

Formulación de hipótesis estadística:

Ho: El machine learning no mejora significativamente el porcentaje de actividades terminadas fuera de tiempo en el Instituto Nacional de Estadística e informática, Lima -2022.

Ha: El machine learning mejora significativamente el porcentaje de actividades terminadas fuera de tiempo en el Instituto Nacional de Estadística e informática, Lima - 2022.

El resultado de la prueba de normalidad del indicador eficacia porcentaje de actividades terminadas fuera de tiempo no es normal se utilizó la prueba de Wilcoxon. También en la tabla 12 se muestra el valor Z que fue de -8.508, el cual se ubica en la zona de rechazo de la hipótesis nula (Ho) y el valor de significancia fue 0.000, el cual es menor del valor del error 0.05, por lo tanto, se rechaza la hipótesis nula (Ho) y se acepta la hipótesis alterna (Ha).

Tabla 12

Medidas descriptivas del indicador

Prueba de rangos con signo de Wilcoxon

		Rangos		
		N	Rango promedio	Suma de rangos
I2Postest - I2Pretest	Rangos negativos	96 ^a	48,50	4656,00
	Rangos positivos	0 ^b	,00	,00
	Empates	0 ^c		
	Total	96		

a. I2Postest < I2Pretest

b. I2Postest > I2Pretest

c. I2Postest = I2Pretest

Estadísticos de prueba^a

	I2Postest - I2Pretest
Z	-8,508 ^b
Sig. asintótica(bilateral)	,000

a. Prueba de rangos con signo de Wilcoxon

b. Se basa en rangos positivos.

Fuente: Elaboración propia

Prueba de Hipótesis específica 3: Eficiencia de colaboradores

Formulación de hipótesis estadística:

Ho: El machine learning no mejora significativamente la eficiencia de colaboradores en el Instituto Nacional de Estadística e informática, Lima -2022.

Ha: El machine learning mejora significativamente la eficiencia de colaboradores en el Instituto Nacional de Estadística e informática, Lima - 2022.

El resultado de la prueba de normalidad del indicador Eficiencia de colaboradores es normal se utilizó la prueba de T Student. En la tabla 13, se observó que la eficiencia de colaboradores tuvo como resultado para la prueba T una valoración de -28.440, este valor se ubica en la zona de rechazo de la hipótesis nula, con un valor p (0.000) que no excede al margen de error de 5% por lo tanto, se rechaza la hipótesis nula (Ho) y se acepta la hipótesis alterna (Ha).

Tabla 13
Medidas descriptivas del indicador

➔ **Prueba T**

Estadísticas de muestras emparejadas

		Media	N	Desv. Desviación	Desv. Error promedio
Par 1	I3Pretest	,5007	96	,08584	,00876
	I3Postest	,8299	96	,07786	,00795

Correlaciones de muestras emparejadas

		N	Correlación	Sig.
Par 1	I3Pretest & I3Postest	96	,043	,680

Prueba de muestras emparejadas

		Diferencias emparejadas							
		Media	Desv. Desviación	Desv. Error promedio	95% de intervalo de confianza de la diferencia		t	gl	Sig. (bilateral)
					Inferior	Superior			
Par 1	I3Pretest - I3Postest	-,32917	,11340	,01157	-,35214	-,30619	-28,440	95	,000

Fuente: Elaboración propia

V. DISCUSIÓN

En la investigación se realizó un análisis del control de proyectos en el INEI, para ello se dimensiono la variable dependiente en tres indicadores que son: eficacia, eficiencia y desempeño. Con la implementación de machine learning se consiguió mejorar el tiempo para tomar decisiones. Se analizó una muestra de 96 proyectos y se recolecto información en fichas de registro. La investigación se realizó mediante un diseño no experimental de corte longitudinal, para esto se usó dos muestras en distintos tiempos que son pretest y postest.

La hipótesis general que dice: machine learning aumenta significativamente el control de proyectos de TI en el INEI concuerda con la investigación de (Caro y otros, 2021) de enfoque cuantitativo, cuyo objetivo principal es resolver las malas prácticas en gestión de proyectos. La información para elaborar el modelo predictivo la obtuvo mediante la observación y de data histórica de proyectos del estado ejecutados en los últimos 10 años. Como resultado general de la investigación es que el modelo entrenado para la evaluación de proyectos tiene una precisión del 79.51%. Se concluye que machine learning predice exitosamente los proyectos a partir de data histórica. La investigación realizada por (Maher, 2021) no concuerda con nuestro trabajo de investigación ya que menciona que machine learning es beneficioso para el seguimiento y control de proyectos ya que el conocimiento que se obtiene del análisis predictivo sirve como referencia para una gestión eficaz y para evitar cuellos de botella referentes a sobre costo , disminución de la eficiencia que demoren la duración del proyecto, pero en los modelos revisados presentan limitaciones ya que no se evalúan todas las dimensiones de los proyectos , entonces el gerente tiene que continuar utilizando su experiencia cuando se trata de liderar equipos de trabajo.

Con respecto a la hipótesis específica 1, machine learning mejora la eficacia de los proyectos de TI en el INEI, Lima ,2022, se aplicó la prueba de Wilcoxon para el pretest, postest y se determinó que no se distribuye normalmente. El estadístico de normalización Z fue de -8.508 y el valor sig. de 0.000, entonces

se acepta la H_a y se rechaza la H_o . Se demostró que la implementación machine learning incremento la eficacia interna de las actividades de un 35.34% a un 87.93% para una muestra de 96 proyectos, lo que representa una mejora del 52.59%. Los resultados concuerdan con el trabajo de (Quipas, 2021) que, mediante el análisis de 50 expedientes registrados en fichas de observación con un muestreo probabilístico, obtuvo la media porcentual del pretest de 42.15% y para el posttest de 81.06% lo que indica una mejora del 38.91% luego de aplicar machine learning. Al realizar el análisis inferencial de expedientes procesados a tiempo, se obtuvo el valor para el grado de significancia de 0.000 obtenido con la prueba no paramétrica de Wilcoxon y para el valor de z de -5.528 que se ubica en la zona de no aceptación de la hipótesis nula. Además, el valor de significancia resultante es inferior al valor asumido de 0.05, entonces, se rechaza la hipótesis nula y se afirma la hipótesis alterna.

Con respecto a la hipótesis específica 2, se aplicó la prueba de Wilcoxon para el pretest, posttest y se determinó que no se distribuye normalmente. El estadístico de normalización Z fue de -8.508 y el valor sig. de 0.000, entonces se afirma la H_a y no se acepta la H_o . Los resultados obtenidos demostraron que la implementación de machine learning logro disminuir la cantidad de actividades terminadas fuera de tiempo de un 71.59% a un 9.27%. Los resultados obtenidos concuerdan con (Silva, 2021) que mediante un análisis descriptivo de una población de 70 transacciones comerciales que se registraron en una ficha de observación se obtuvo una media del pretest de 66.28% y el valor del posttest de 8.56% indicando una mejora del 57.72% de menos solicitudes atendidas después de tiempo .Al realizar el análisis inferencial de solicitudes atendidas fuera de tiempo se obtuvo un valor de significancia para el pretest de 0.016 y posttest de 0.000 , estos valores son inferiores al error asumido de 0.005. Para comprobar la hipótesis se realizó pruebas no paramétricas como la de Wilcoxon y se obtuvo el valor de significancia de 0.000 siendo inferior al valor alfa de 0.005, entonces no se acepta la hipótesis nula. El autor concluye que el índice de solicitudes atendidas fuera de tiempo mejora al implementarse machine learning. Así también (López J. , 2019) en su trabajo de investigación cuantitativa y como

instrumento la observación y ficha de recolección de datos pudo predecir si un cliente de una institución financiera dedicada a las microfinanzas era un buen o mal pagador. Para reducir el riesgo de clientes morosos, se le solicita información económica en detalle. El objetivo principal del trabajo de investigación fue comparar diferentes algoritmos para reportar con exactitud la predicción de incumplimiento de pagos. El autor ha utilizado varios algoritmos de machine learning como: Regresión logística, Random Forest, Perceptron multicapa, XGBoost y Gradient Boosting para pronosticar la falta de pago por parte del cliente. El modelo de machine learning que resulto con mayor exactitud fue el XGBoost con 97.53%. También (Yaranga, 2022) en su investigación de enfoque cuantitativo con una muestra de 120 observaciones, como instrumento de recolección de datos empleo a la ficha de observación para realizar la prueba de hipótesis donde usó la prueba de Wilcoxon que permitió realizar una comparación entre el pretest y posttest que determinó una mejora de los indicadores: 36.9% para personal postulante, 40.3% para contratación y 36.8% para evaluación curricular. Así también (Valle, 2021) en su investigación de enfoque cuantitativo y del tipo aplicada demostró que implementar un sistema de información aminora el tiempo de resolución de documentos del 68.63% a 56.81% para una muestra de 357 documentos procesados.

De los resultados obtenidos de la hipótesis 3, se demostró que machine learning incremento la eficiencia de los colaboradores de un 50.07% a un 82.99%, lo que representa una mejora del 32.92%. Respecto a la media de la eficiencia de colaboradores, se realizó la comprobación de normalidad con la prueba de Kolmogorov, donde se obtuvo el grado de significancia para el pretest (0.200) y posttest (0.200) donde ambas superaron el 5% del margen de error, lo cual indica normalidad, entonces se utilizó pruebas paramétricas como la T-Student y se obtuvo un valor para el grado de significancia de 0.000. También para el valor z se obtuvo el valor -28.440 que se encuentra en la zona de no aceptación de la hipótesis nula, por lo cual se acepta la hipótesis alterna que menciona que machine learning mejora significativamente la eficiencia de los colaboradores. También (Coronel, 2021) en su investigación cuantitativa realizo un posttest y un

pretest teniendo como muestra 300 observaciones donde utilizo a la variable independiente que es machine learning sobre los efectos que ocasiona en la variable dependiente que son los indicadores. Entonces para el proceso de selección del personal administrativo se calculó una media para el pretest de 20.3% y para el posttest 22.3%, lo cual evidencia una mejora del 2%. Entonces se evidencia que al implementar machine learning mejora el indicador. En cuanto al análisis de la normalidad para el posttest y pretest se obtuvo el valor de 0.000 para ambos casos, por lo cual se orientó medidas no paramétricas como la prueba de Wilcoxon y se calculó el valor para el grado de significancia de 0.000 y con este resultado no se acepta la hipótesis nula y se acepta la hipótesis alterna. También (Yamao, 2018) en su investigación de enfoque cuantitativo y diseño de investigación del tipo correlacional-causal, pretende encontrar la relación de rendimiento académico con factores económicos y sociales de ingresantes a la universidad para predecir el rendimiento de los alumnos. El autor utilizó una muestra probabilística con un intervalo de confianza de 95%. También para realizar predicciones se utilizó tres técnicas: support vector machines, regresión lineal, árbol de decisiones y donde la mejor predicción la obtuvo el árbol de decisiones que fue del 82.87%. Al implementar el algoritmo permitió detectar a los alumnos ingresantes que podrían presentar problemas en sus estudios académicos. En su estudio (Li y otros, 2019) propone un algoritmo de machine learning para resolver problemas de decisión inteligente con el objetivo de predecir la eficiencia en la innovación de empresas. El autor para elaborar el modelo predictivo se basó en la observación y recopilación de información de 10 años donde analizó la eficiencia de 10 mil empresas. El modelo desarrollado permite a los gerentes predecir el desempeño de la innovación en la empresa, además incrementa la eficiencia a largo respecto a otras organizaciones. También el autor ha empleado 4 algoritmos de machine learning como: regresión logística, árbol de decisiones, bosques aleatorios, redes neuronales y XGBOOST. Al analizar los 4 algoritmos se obtiene mejor precisión con XGBOOST que fue de 77%, entonces se llega a la conclusión que machine learning y uso de información histórica de expedientes predicen la evaluación y comportamiento de proyectos permitiendo mejorar la toma de decisiones.

VI. CONCLUSIONES

Primera: Los resultados de la investigación concluye que machine learning aumenta significativamente el control de proyectos de TI en el INEI, Lima ,2022. Después de implementar machine learning el control de proyectos mejoró de un 52.39% a un 60.01%. El valor de significancia obtenido fue 0.000, entonces es menor del valor del error 0.05, por lo tanto, se niega la hipótesis nula (H_0) y se acepta la hipótesis alterna (H_a).

Segunda: Los resultados obtenidos evidenciaron en la hipótesis específica 1, que mediante el indicador eficacia interna de las actividades, mejoró de un 35.34% a un 87.93%. En el indicador la mejora se dio en el incremento del 52.59%. También se obtuvo un valor estadístico para Z de -8.508 con un valor de significancia de 0.000 que determina la no aceptación de la hipótesis nula y la afirmación de la hipótesis alterna.

Tercera: Los resultados obtenidos evidenciaron en la hipótesis específica 2, que mediante el indicador actividades terminadas fuera de tiempo, se puede concluir que el índice de actividades terminadas fuera de tiempo progreso de un 71.59% a un 9.27%. También se obtuvo un valor estadístico para Z de -8.508 con un valor de significancia de 0.000 que determina la no aceptación de la hipótesis nula y la afirmación de la hipótesis alterna.

Cuarta: Los resultados obtenidos evidenciaron en la hipótesis específica 3, que mediante el indicador Eficiencia de colaboradores aumentó de un 50.07% a un 82.99%. También se obtuvo un valor estadístico para Z de -28.440 con un valor de significancia de 0.000 que determina la no aceptación de la hipótesis nula y la afirmación de la hipótesis alterna.

VII. RECOMENDACIONES

- Primera: Para la continuidad de los resultados positivos de las tres hipótesis específicas realizadas en el INEI, después de implementar machine learning, se recomienda mayor explotación de los datos y la creación de nuevos indicadores que mejoren el control de los proyectos.
- Segunda: Para la hipótesis 1, se recomienda que el director técnico priorice que actividades son urgentes y tener en cuenta el tiempo de demora que puede retrasar a las actividades importantes
- Tercera: Respecto a la hipótesis 2, se sugiere al supervisor que automatice la tarea de recolectar el avance de cada empleado mediante modelos predictivos, pues puede anticipar las tareas que no van a poder terminarse a tiempo.
- Cuarta: Respecto a la hipótesis 3, se sugiere al supervisor considere los resultados obtenidos para la eficiencia de los colaboradores para reasignar recursos humanos para cada tipo de tareas, también se tiene que crear nuevos hitos para el control del personal

Referencias

- A.S., G., & Edayadiyil, J. (2022). Automated progress monitoring of construction projects using Machine learning and image processing approach. ScienceDirect. <https://doi.org/https://doi.org/10.1016/j.matpr.2022.03.137>
- Akbar, T., Amirhessam, T., Banzhaf, W., & Gandomi, A. (2021). Evolutionary Machine Learning: A Survey. ACM Computing Surveys., 1-35. <https://doi.org/10.1145/3467477>
- Álvarez Ochoa, J. Ó. (2021). Ingeniería Industrial, edición especial/2022, ISSLa inteligencia artificial en la gestión de proyectos de inversión pública del Ministerio de Vivienda, Construcción y Saneamiento. Revista ULima, 97-121. <https://doi.org/https://doi.org/10.26439/ing.ind2022.n.5802>
- Alzubi, J., Nayyar, A., & Kumar, A. (2018). Machine Learning from Theory to Algorithms: An Overview. IOP Publishing. <https://doi.org/doi:10.1088/1742-6596/1142/1/012012>
- Andrade, A., Del Rio, C., & Alvear, D. (2019). Estudio de Tiempos y Movimientos para Incrementar la Eficiencia en una Empresa de Producción de Calzado. Scielo, 30(3). <https://doi.org/http://dx.doi.org/10.4067/S0718-07642019000300083>
- Banco Interamericano de Desarrollo. (2018). HERRAMIENTAS Y TÉCNICAS PARA LA GESTIÓN DE PROYECTOS DE DESARROLLO PM4R (4 ed.). Banco Interamericano de Desarrollo. https://indesvirtual.iadb.org/file.php/1/PM4R/Guia%20de%20Aprendizaje%20PMA%20SPA.pdf?fbclid=IwAR0_17MRzWGU-xgLTa1HregQQYcDu4V8vVnAga7GbhPdR2dJ0QbezaNZ-ig
- Caro, J., Cruz, C., & Nova, L. (2021). DISEÑO DE UN MODELO DE EVALUACIÓN DE PROYECTOS A PARTIR DE 2 HERRAMIENTAS DE MACHINE LEARNING O APRENDIZAJE AUTOMATIZADO [Tesis maestría, Escuela Colombiana de Ingeniería Julio Garavito]. Repositorio

institucional, Colombia.

<https://repositorio.escuelaing.edu.co/bitstream/handle/001/1444/Nova%20Barreto%2C%20Laura%20Tatiana-2021.pdf?sequence=3&isAllowed=y>

Chio, C., & Freeman, D. (2018). *Machine Learning and Security: Protecting Systems with Data and Algorithms* (1 ed.). (C. Allen, Ed.) Estados Unidos: O'Reilly.

https://books.google.com.pe/books?hl=es&lr=&id=mSJJDwAAQBAJ&oi=fnd&pg=PR4&dq=chio+machine+learning&ots=ryL8Dw4N8h&sig=Sof7Bm_Tlf6L6a-NLt-2ST9ZvcY#v=onepage&q=chio%20machine%20learning&f=false

Coronel, E. (2021). *Machine Learning en la mejora del proceso de selección del personal administrativo de la Corte Superior de Justicia de Lima, 2020*[Tesis de maestría, Universidad Cesar Vallejo]. Repositorio institucional.

https://repositorio.ucv.edu.pe/bitstream/handle/20.500.12692/61903/Coronel_CEG-SD.pdf?sequence=1&isAllowed=y

Duque, J. d., & Supo, D. (2021). MEJORA DEL DESEMPEÑO DE UNA EMPRESA CONTRATISTA Y DE SERVICIOS GENERALES APLICANDO LA METODOLOGÍA DE LA GUÍA PMBOK EN LA GESTIÓN DE PROYECTOS. *Revista Ingeniería: Ciencia, Tecnología e Innovación*, 8.

<https://revistas.uss.edu.pe/index.php/ING/article/view/1538/2242>

Garbero, A., Carneiro, B., & Resce, G. (2021). Harnessing the power of machine learning analytics to understand food systems dynamics across development projects. *ScienceDirect*, 172.

<https://doi.org/https://doi.org/10.1016/j.techfore.2021.121012>

García, J., Cazallo, A., Barragan, C., Mercado, M., Olarte, L., & Meza, V. (2019). Indicadores de Eficacia y Eficiencia en la gestión de procura de materiales en empresas del sector construcción del Departamento del Atlántico, Colombia. *Espacios*, 40(22).

<https://www.revistaespacios.com/a19v40n22/a19v40n22p16.pdf>

- Gonzalez, C., Ril, E., Gonzalez, H., Milian, V., & Camejo, J. (2019). Descubrimiento de conocimientos en los comentarios que realizan los usuarios en un sistema de noticias digital. *Revista Cubana de Ciencias Informáticas*, 13(4), 56-66.
<https://www.redalyc.org/journal/3783/378365914005/html/>
- Hasan, S., Piñero, P., García, R., Sadeq, G., & Alvarado, L. (2020). Nuevo modelo de toma de decisiones basado en mapas cognitivos neutrosóficos triangulares para la evaluación de proyectos. *Revista Investigacion Operacional*, 41(5), 768-779.
https://www.researchgate.net/publication/346534430_NUEVO_MODELO_DE_TOMA_DE_DECISIONES_BASADO_EN_MAPAS_COGNITIVOS_NEUTROSOFICOS_TRIANGULARES_PARA_LA_EVALUACION_DE_PROYECTOS
- Henríquez, E. (2021). Aplicación de la tecnología QR en dispositivos móviles para mejorar el control de ingreso del personal en los fondos de Chao de la Empresa Tal S.A. [Tesis de maestría, Universidad Cesar Vallejo]. Repositorio institucional, Perú.
<https://hdl.handle.net/20.500.12692/55948>
- Hernández, R., Fernández, C., & Baptista, M. (2014). *Metodología de la investigación* (6 ed.). México: McGRAW-HILL.
- Hernández, S., & Duana, D. (2020). Técnicas e instrumentos de recolección de datos. *Boletín Científico de las Ciencias Económico Administrativas del ICEA*, 9(17), 51-53.
<https://repository.uaeh.edu.mx/revistas/index.php/icea/issue/archive>
- Hurwitz, J., & Kirsch, D. (2018). *Machine Learning for dummies*. (R. Mengle, Ed.) John Wiley & Sons, Inc.
<https://www.ibm.com/downloads/cas/GB8ZMQZ3>
- Instituto Nacional de Estadística e Informática. (2017). *PLAN ESTRATÉGICO DE GOBIERNO ELECTRÓNICO (PEGE) DEL INSTITUTO NACIONAL DE ESTADÍSTICA E INFORMÁTICA (INEI) 2017 – 2019*.

https://www.inei.gob.pe/media/pte/PEGE/Plan_Estrategico_GobElect-PEGE_VF.PDF

- Kashyap, P. (2017). Machine Learning Algorithms and Their Relationship with Modern Technologies. OUCI, 91-136.
https://doi.org/https://doi.org/10.1007/978-1-4842-2988-0_3
- Krishna, M., Amirali, B., Swathi, A., & Madhavi, V. (2021). High performance machine learning and data science based implementation using Weka. ScienceDirect. <https://doi.org/https://doi.org/10.1016/j.matpr.2021.01.470>
- Kumar, A., Ganapathysubramanian, B., Sarkar, S., & Singh, A. (2018). Deep Learning for Plant Stress Phenotyping: Trends and Future Perspectives. ScienceDirect, 23, 883-898.
<https://doi.org/https://doi.org/10.1016/j.tplants.2018.07.004>
- León, C. (2020). Modelo de solución de Business Intelligence y Machine Learning para el monitoreo y control de calidad de la medición del consumo de agua en el Centro de Servicios Breña [Tesis de maestría, Universidad ESAN]. Repositorio institucional.
https://repositorio.esan.edu.pe/bitstream/handle/20.500.12640/2142/2020_MADTI_13-1_07_T.pdf?sequence=1&isAllowed=y
- Li, Y., Yang, L., Yang, B., Wang, N., & Wu, T. (2019). Application of interpretable machine learning models for the. ELSEVIER, 273-283.
<https://doi.org/https://doi.org/10.1016/j.neucom.2018.12.012>
- López, J. (2019). COMPARACIÓN DE MODELOS DE APRENDIZAJE DE MÁQUINA EN LA PREDICCIÓN DEL INCUMPLIMIENTO DE PAGO EN EL SECTOR DE LAS MICROFINANZAS[Tesis de maestría, Pontificia Universidad Católica del Perú]. Repositorio institucional.
https://tesis.pucp.edu.pe/repositorio/bitstream/handle/20.500.12404/19546/LOPEZ_MALCA_JIAN_COMPARACION_MODELOS.pdf?sequence=1&isAllowed=y
- López, M. Á., Vázquez, P., & Quiñoá, L. (2022). An approach to employees' job performance through work environmental variables and leadership

- behaviours. ScienceDirect, 140, 361-369.
<https://doi.org/https://doi.org/10.1016/j.jbusres.2021.11.006>
- Lucero, D. (2022). Herramienta gráfica para la automatización de historias de usuario desde un modelo orientado a objetivos con gestión de trazabilidad [Tesis de maestría: Universidad Mayor de San Marcos]. Repositorio institucional, Perú.
https://cybertesis.unmsm.edu.pe/bitstream/handle/20.500.12672/17620/Lucero_gd.pdf?sequence=1&isAllowed=y
- Maher, M. (2021). Applications of Artificial Intelligence in Project Management [Tesis de maestría, Politecnico Di Torino]. Repositorio institucional, Italia.
https://webthesis.biblio.polito.it/view/creators/Mohamed=3AKhaled_Maher_Abdalla=3A=3A.default.html
- Mero, J. (2018). Empresa, administración y proceso administrativo. Revista Científica ciencias económicas y empresariales, 3(8).
<https://doi.org/https://doi.org/10.23857/fipcaec.v3i8.59>
- Mis, G., & Sankaranarayanan, S. (2021). Prediction of Risk Percentage in Software Projects by Training Machine Learning Classifiers. ScienceDirect, 94.
<https://doi.org/https://doi.org/10.1016/j.compeleceng.2021.107362>
- Mora, E. (2018). Algoritmo Machine Learnig en los sistemas de filtrado caso práctico SPAM de google en las cuentas de correo Institucionales de la FAFI. Pro Sciences, 2(12). <https://doi.org/https://doi.org/10.29018/issn.2588-1000vol2iss12.2018pp17-23>
- Musbah, H., Ali, G., Aly, H., & Little, T. (2022). Energy management using multi-criteria decision making and machine learning classification algorithms for intelligent system. ScienceDirect, 203.
<https://doi.org/https://doi.org/10.1016/j.epsr.2021.107645>
- Najah, M., Mohamed, M., Rahim, A., Ismail, R., Kok, L., Sufyian, M., Natiq, H., & Happala, H. (2021). Software Project Management Using Machine

- Learning Technique—A Review. (V. Conforti, Ed.) MDPI.
<https://doi.org/https://doi.org/10.3390/app111115183>
- Nwanyanwu, L. (2018). Accountants' Ethics and Fraud Control in Nigeria: The Emergence of a Fraud Control Model. *Jafas*, 130-150.
https://jafas.org/articles/2018-4-1/8_Ethics_FULL_TEXT.pdf
- Otzen, T., & Manterola, C. (2017). Técnicas de Muestreo sobre una Población a Estudio. (M. d. Sol, Ed.) *International Journal of Morphology*, 227-232.
<https://scielo.conicyt.cl/pdf/ijmorphol/v35n1/art37.pdf>
- Pan, Y., & Stark, R. (2022). An interpretable machine learning approach for engineering change management decision support in automotive industry. *ScienceDirect*, 138.
<https://doi.org/https://doi.org/10.1016/j.compind.2022.103633>
- Pellerin, R., & Perrier, N. (2019). A review of methods, techniques and tools for project planning and control. *Taylor&Francis Online*, 57.
<https://doi.org/https://doi.org/10.1080/00207543.2018.1524168>
- Project Management Institute. (2018). El éxito en tiempos de disrupción.
https://www.pmi.org/-/media/pmi/documents/public/pdf/learning/thought-leadership/pulse/pulse-of-the-profession-2018.pdf?sc_lang_temp=es-ES
- protiviti. (2019). Competing in the cognitive age: how companies will transform their businesses and drive value through advanced AI.
https://www.protiviti.com/sites/default/files/united_states/insights/ai-ml-global-study-protiviti.pdf
- Quipas, M. (2021). Machine Learning en la mejora de la gestión del conocimiento en el Consejo Nacional de Ciencia y Tecnología e Innovación Tecnológica - CONCYTEC, Lima 2021 [Tesis de maestría, Universidad Cesar Vallejo]. Repositorio institucional.
<https://hdl.handle.net/20.500.12692/71721>
- Rudas, L. (2017). MODELO DE GESTION DE RIESGOS PARA PROYECTOS DE DESARROLLO TECNOLOGICO [Tesis de maestría, Ciateq]. Repositorio institucional, México.

<https://ciateq.repositorioinstitucional.mx/jspui/bitstream/1020/86/1/RudasTayoLeidyP%20MDGPI%202017.pdf>

Sánchez, J. (2021). Aplicación móvil con geolocalización para mejorar el control de siniestros ocurridos por accidentes de tránsito – Afocat Chimbote Región Ancash [Tesis de maestría , Universidad César Vallejo]. Repositorio institucional, Perú.
<https://repositorio.ucv.edu.pe/handle/20.500.12692/56051>

Sazili, M., Anuar, S., Hussin, N., Ibrahim, Z., Ali, J., & Wan, S. (2019). FACTORS INFLUENCING INFORMATION TECHNOLOGY PROJECT MANAGEMENT SUCCESS IN THE FINANCIAL INDUSTRY. JATIT, 97(10). <http://www.jatit.org/volumes/Vol97No10/11Vol97No10.pdf>

Silva, J. (2021). Machine Learning en la mejora del proceso de operaciones comerciales en la empresa Redondos, Lima 2020 [Tesis de maestría, Universidad Cesar Vallejo]. Repositorio institucional.
<https://repositorio.ucv.edu.pe/handle/20.500.12692/63573#:~:text=La%20implementaci%C3%B3n%20del%20machine%20learning,del%20proceso%20de%20operaciones%20comerciales.>

Sorooshian, S., & Yin, S. (2020). Literature Review: Critical Risk Factors. ProQuest, 21(175), 157-161.
<https://www.proquest.com/docview/2381628490/fulltextPDF/1610147983C4697PQ/1?accountid=37408>

Srikrishna, C., & Shivendra, P. (2022). Artificial intelligence in marketing: A systematic literature review. International Journal of Market Research, 64, 38-68. <https://doi.org/10.1177/14707853211018428>

Valbuena, R. (2021). Inteligencia artificial: Investigación científica avanzada centrada en datos (Primera edición ed.). Venezuela: Cencal Press.
https://books.google.com.pe/books/about/Inteligencia_Artificial.html?id=SoMTEAAAQBAJ&redir_esc=y

Valle, S. (2021). Implementación de un sistema de información para el proceso documentario en el Instituto de Radio y Televisión del Perú 2021[Tesis

de maestría , Universidad César Vallejo]. Repositorio institucional, Perú.
<https://doi.org/https://orcid.org/0000-0001-5207-9353>

Vigo, V., Vigil, S., Sánchez, M., & Medianero, D. (2018). Manual de diseño de proyectos de Desarrollo Sostenible (2 ed.). (B. N. Perú, Ed.) Perú.
<https://www.losandes.org.pe/libros/LIBRO-DISENO-DE-PROYECTOS-ALAC.pdf>

Vrchota, J., & Rehor, P. (2021). PROJECT MANAGEMENT IN MANUFACTURING ENTERPRISES. Serbian Journal of Management, 16, 341-353. <https://doi.org/10.5937/sjm16-28044>

Yamao, E. (2018). PREDICCIÓN DEL RENDIMIENTO ACADÉMICO MEDIANTE MINERÍA DE DATOS EN ESTUDIANTES DEL PRIMER CICLO DE LA ESCUELA PROFESIONAL DE INGENIERÍA DE COMPUTACIÓN Y SISTEMAS, UNIVERSIDAD DE SAN MARTÍN DE PORRES, LIMA-PERÚ[Tesis de maestría,USMP]. Repositorio institucional.
https://repositorio.usmp.edu.pe/bitstream/handle/20.500.12727/3555/yamao_e.pdf?sequence=3&isAllowed=y

Yaranga, I. (2022). Machine Learning en la mejora del Proceso de Selección del Personal Docente en una Universidad Nacional, Lima 2021[Tesis de maestría,Universidad Cesar Vallejo]. Repositorio institucional, Perú.
<https://repositorio.ucv.edu.pe/handle/20.500.12692/85185>

ANEXOS

Anexo 1: Matriz de Operacionalización de Variables

VARIABLES DE ESTUDIO	DEFINICIÓN CONCEPTUAL	DEFINICIÓN OPERACIONAL	DIMENSIÓN	INDICADORES	ESCALA DE MEDICIÓN
Machine learning	Machine learning o aprendizaje automático es un subconjunto de técnicas de inteligencia artificial que utiliza algoritmos para extraer patrones mediante métodos estadísticos, matemáticos para optimizar y descubrir conocimientos. (Akbar y otros, 2021)	Se refiere al uso de data histórica para diseñar algoritmos predictivos para pronosticar resultados futuros			
Control de proyectos	El control de proyectos es una de las funciones críticas en la gestión de proyectos, entonces, implica ejecutar varias decisiones para programar los proyectos, informar del estado del proyecto comparándolo con el plan inicial,	Para el control de proyectos implica un conjunto de decisiones que permiten anticiparnos a situaciones que pueden salir del alcance planificado antes de empezar a realizar el proyecto	Eficacia	Eficacia interna de las actividades	Razón
			Desempeño	Actividades terminadas fuera de tiempo	Razón
			Eficiencia	Eficiencia de colaboradores	Razón

	detección de situaciones fuera de control, tomar decisiones apropiadas y analizar situaciones. (Pellerin & Perrier, 2019)				
--	---	--	--	--	--

Anexo 2: Instrumento de recolección de datos

Ficha de observación N01 – Indicador eficacia interna de las actividades

Este indicador se utiliza para medir la disposición de la organización para lograr la satisfacción del cliente y el logro de objetivos.

Ficha de Observación de eficacia interna de las actividades					
Investigador		Iván Wilber Paredes Reyes			
Proceso Observado		Control de proyecto			
OBSERVACIÓN Y REGISTRO					
Nro	Producto	Fecha	Cantidad de actividades realizadas	Cantidad de actividades programadas	EIA = (Cantidad de actividades realizadas / Cantidad de actividades programadas)
1					
2					
3					
4					
N					



Dr. Marlon Acuña Benites

DNI: 42097456

Ing. de Sistemas / Investigador

Ficha de observación N02 – Indicador actividades terminadas fuera de tiempo

Este indicador se relaciona con la habilidad que tiene la organización para cumplir las metas, la misión y actividades para controlar las necesidades del interesado.

Ficha de Observación de actividades terminadas fuera de tiempo					
Investigador		Iván Wilber Paredes Reyes			
Proceso Observado		Control de proyecto			
OBSERVACIÓN Y REGISTRO					
Nro	Producto	Fecha	Tareas que se terminaron a tiempo	Número total de tareas	ATFT =(Tareas que se terminaron a tiempo / Total de tareas) *100
1					
2					
3					
4					
N					



Dr. Marlon Acuña Benites
DNI: 42097456
Ing. de Sistemas / Investigador

Ficha de observación N03 – Indicador eficiencia de colaboradores

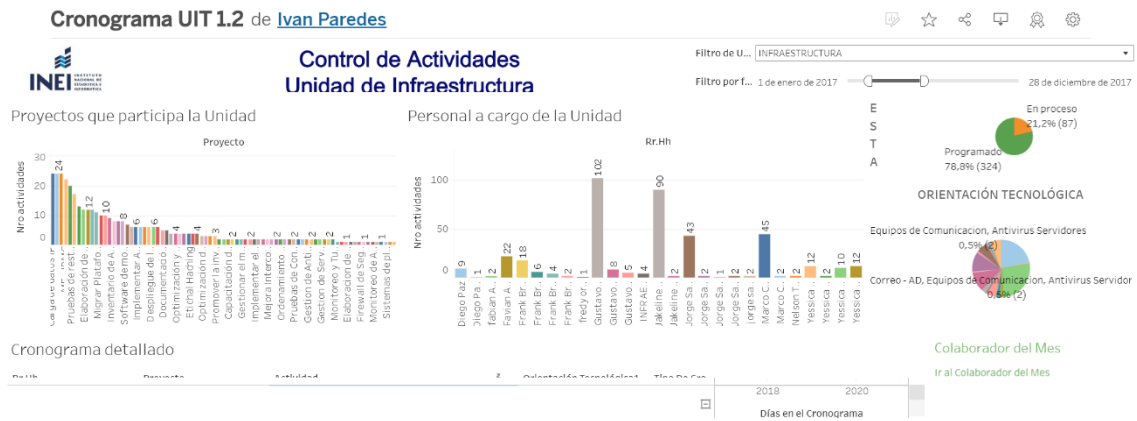
Este indicador se utiliza para medir el cumplimiento de los proyectos y el uso de los recursos mediante una relación entre recursos utilizados y programados. También sirve para medir la calidad de recursos al transformarse en producto o servicio.

Ficha de Observación de eficiencia de colaboradores					
Investigador			Iván Wilber Paredes Reyes		
Proceso Observado			Control de proyecto		
OBSERVACIÓN Y REGISTRO					
Nro	Producto	Fecha	Total de proyectos resueltos	Cantidad total de proyectos	EDC =(Total de proyectos resueltos / Cantidad total de proyectos) *100
1					
2					
3					
4					
N					



Dr. Marlon Acuña Benites
DNI: 42097456
Ing. de Sistemas / Investigador

Anexo 3: Imágenes de sistema



Cronograma UIT de Ivan Paredes

Hoja 6

Proyecto





Declaratoria de Autenticidad del Asesor

Yo, ACUÑA BENITES MARLON FRANK, docente de la ESCUELA DE POSGRADO MAESTRÍA EN INGENIERÍA DE SISTEMAS CON MENCIÓN EN TECNOLOGÍAS DE LA INFORMACIÓN de la UNIVERSIDAD CÉSAR VALLEJO SAC - LIMA NORTE, asesor de Tesis titulada: "Machine Learning para el Control de Proyectos de TI en el Instituto Nacional de Estadística e Informática, Lima, 2022", cuyo autor es PAREDES REYES IVAN WILBER, constato que la investigación cumple con el índice de similitud establecido, y verificable en el reporte de originalidad del programa Turnitin, el cual ha sido realizado sin filtros, ni exclusiones.

He revisado dicho reporte y concluyo que cada una de las coincidencias detectadas no constituyen plagio. A mi leal saber y entender la Tesis cumple con todas las normas para el uso de citas y referencias establecidas por la Universidad César Vallejo.

En tal sentido, asumo la responsabilidad que corresponda ante cualquier falsedad, ocultamiento u omisión tanto de los documentos como de información aportada, por lo cual me someto a lo dispuesto en las normas académicas vigentes de la Universidad César Vallejo.

LIMA, 08 de Agosto del 2022

Apellidos y Nombres del Asesor:	Firma
ACUÑA BENITES MARLON FRANK DNI: 42097456 ORCID 0001-5207-9353	Firmado digitalmente por: MACUNABE el 08-08- 2022 10:52:54

Código documento Trilce: TRI - 0402332