**UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS**

FACULTAD DE INGENIERÍA DE SISTEMAS E INFORMATICA 

**CURSO:**

[AUTOMATIZACIÓN Y CONTROL DE SOFTWARE](https://classroom.google.com/u/1/c/NzA3Njc4MTg0Mzk2)

**TEMA:**

----------

**PROFESOR:**

YESSICA ROSAS CUEVA

**ESTUDIANTES:**

* Felix Huayhua, Axel Patrick
* Fernandez Camacho, Geomar Willy
* C
* Saavedra Monterrey Max Bruno
* E
* Sanchez Saldaña, Sergio Antonio

**Lima - Perú**

**2024**

Resumen

El resumen ha de tener una longitud de un párrafo de entre 150 y 250 palabras sin sangría. Los títulos de sección, como la palabra Resumen anterior, no se consideran títulos, por lo que no se usa formato de título en negrita. En su lugar, use el estilo Título de sección. Este estilo inicia automáticamente la sección en una nueva página, por lo que no es necesario que agregue saltos de página. (Para ver el documento con la paginación, seleccione la pestaña Vista y haga clic en Vista de lectura). Tenga en cuenta que todos los estilos de texto de esta plantilla están disponibles en la pestaña Inicio de la cinta, en la galería de estilos.

Palabras clave: Agregue aquí las palabras clave. Para reemplazar este (o cualquier) texto de sugerencia por el suyo, selecciónelo y comience a escribir. No incluya espacios a la izquierda o a la derecha de los caracteres de la selección.

Capítulo 1: Introducción

* 1. Antecedentes del Problema
  2. Formulación del problema
  3. Justificación
  4. Objetivos
     1. Objetivo general
     2. Objetivos específicos
  5. Organización del Artículo

Capítulo 2: Marco teórico

Péndulo Invertido

Algoritmo

Aprendizaje por Refuerzo

DQ Network

Q-Learning

SARSA

Capítulo 3: Estado del arte

## Artículo 1

**Título:** Aprendizaje por refuerzo profundo difuso guiado por conocimiento

**Aporte:** El artículo propone un nuevo método de aprendizaje por refuerzo profundo, llamado KFDQN, que utiliza sistemas difusos para integrar conocimiento previo en el proceso de aprendizaje. Esto mejora la eficiencia del aprendizaje, la estabilidad del entrenamiento y el rendimiento general del modelo. KFDQN integra el conocimiento en la selección de acciones y en la definición de objetivos de aprendizaje. Se valida el algoritmo mediante simulaciones y experimentos con robots móviles, mostrando una mejora significativa en eficiencia y rendimiento. El estudio destaca la importancia de combinar el aprendizaje automático con el conocimiento humano para lograr mejores resultados.

**Proceso:** El estudio se basa en la combinación de aprendizaje por refuerzo profundo (DQN) con un sistema difuso para representar el conocimiento. El método HYAS (Hybrid Action Selection Strategy) se usa para integrar el conocimiento en la selección de acciones, reduciendo la aleatoriedad. HYL (Hybrid Learning Method) combina conocimiento y aprendizaje automático en el objetivo de aprendizaje. Se utilizan tareas en OpenAI Gym (CartPole, MountainCar) y en robots móviles (llegar a un objetivo y evitar obstáculos) para validar el método. Se evalúa el rendimiento en términos de eficiencia de aprendizaje, estabilidad y rendimiento del modelo.

**Resultados:** Los resultados experimentales demuestran una mejora significativa en la eficiencia del aprendizaje (28.6%), el rendimiento (19.56%) y la estabilidad del entrenamiento de KFDQN en comparación con los algoritmos base. KFDQN mostró una mejor eficiencia de aprendizaje en tareas de planificación a corto y largo plazo. En robots móviles, KFDQN logró mejor rendimiento en tareas de alcanzar un objetivo y esquivar obstáculos. KFDQN superó a algoritmos base en estabilidad y tiempos de ejecución en ambas tareas. El método propuesto destaca la eficacia de la integración del conocimiento previo en el aprendizaje por refuerzo.

## Artículo 2

**Título:** Control de péndulo invertido: un estudio comparativo entre el control convencional y el aprendizaje por refuerzo

**Aporte:** Este artículo se aborda la comparación entre métodos de control convencionales (compensador basado en locus de raíces, compensador de estado y control proporcional-derivativo, PD) y un enfoque basado en aprendizaje por refuerzo (Proximal Policy Optimization, PPO) para controlar un sistema de péndulo invertido. Su contribución principal radica en evidenciar que, bajo un entrenamiento adecuado, el aprendizaje por refuerzo puede lograr un rendimiento comparable a los métodos tradicionales sin requerir conocimiento previo del modelo del sistema. Esto resulta valioso en sistemas donde el modelado es complejo o inaccesible.

**Proceso:** El procedimiento descrito en el artículo se dividió en tres fases principales. En el primer apartado, se trató el proceso de modelado y formulación del sistema de péndulo invertido. Mediante los principios de las leyes de Newton, se ha desarrollado un modelo dinámico que explica las relaciones entre el vehículo y el péndulo. Posteriormente, se procedió a linealizar este modelo no lineal con el fin de simplificar la implementación de técnicas de control convencionales. Para la metodología basada en el aprendizaje por refuerzo, se creó una función de recompensa que tuvo en cuenta la ubicación y la inclinación del péndulo, con el objetivo de mantenerlo en una posición estable.  
En la fase dos, se llevaron a cabo la implementación de los algoritmos de control. Se han elegido tres métodos de control convencionales: el compensador basado en el lugar de las raíces, el compensador de estado y el control proporcional-derivativo (PD). Además, se empleó el algoritmo de optimización de políticas proximales (PPO) como un ejemplo de los enfoques de aprendizaje por refuerzo. El último modelo fue entrenado con 250,000 episodios en un entorno de simulación modificado de OpenAI Gym. Este entorno es compatible con acciones continuas y proporciona recompensas basadas en el comportamiento del sistema.  
Por último, en la tercera fase, se procedió a evaluar los algoritmos a través de pruebas realizadas en un modelo de simulación del péndulo invertido. Se examinaron las respuestas temporales, abarcando el tiempo máximo, el tiempo de estabilización y el exceso de respuesta, además de las respuestas en equilibrio, teniendo en cuenta el error en equilibrio y la energía total consumida.

**Resultados:** El examen de los resultados indicó que el compensador que se basa en el locus de raíces tuvo el rendimiento más sobresaliente en general, presentando un tiempo de estabilización rápido y un error en estado estacionario mínimo. Este procedimiento consiguió estabilizar el péndulo de forma eficaz, destacándose sobre otros enfoques en diversos aspectos fundamentales. El algoritmo PPO ha mostrado ser competitivo, especialmente en cuanto a eficiencia energética, al requerir menos energía para estabilizar el sistema en comparación con las técnicas convencionales. No obstante, el rendimiento durante el periodo de estabilización y el momento de mayor actividad fue deficiente, esto se debió a las restricciones de la función de recompensa empleada, la cual no incluyó de manera explícita criterios de rendimiento temporal. Por otra parte, los métodos de compensación de estado y control proporcional-derivativo mostraron un rendimiento más restringido. Específicamente, el control proporcional-derivativo presentó oscilaciones continuas y no logró estabilizar adecuadamente el péndulo en el tiempo asignado para las pruebas.

## Artículo 3

## Artículo 4

**Titulo:** Control basado en Q-learning para balanceo y equilibrio de péndulo invertido

**Aporte:** La investigación vincula la teoría de control clásica con técnicas

De aprendizaje por refuerzo, contribuyendo para demostrar versatilidad y potencial del aprendizaje automático en tareas de control. Se identificó que el factor de descuento óptimo influye significativamente en el resultado del aprendizaje y en la estabilidad del sistema.

**Proceso:** Para la simulación del modelo se empleó una representación precisa de la dinámica física la cual rigue al sistema, combinada con un algoritmo de aprendizaje por refuerzo para el control de movimientos del carro, las fuerzas actuantes son la fuerza aplicada la fuerza gravitacional del pendulo y la inercia del pendulo. Se establecio un limite en cada lado con una distancia igual desde la posición inicial del carro, lo que se busca es que el angulo en el que oscila el pendulo este dentro de los 20 grados. Con relación al estado, el estado siguiente responde al estado actual y a la acción realizada, los parámetros actualizados derivan de ecuaciones físicas simplificadas. Se hicieron usos de 3 tasas de aprendizajes (0.5,0.7 y 0.9) y 3 factores de descuentos (0.1,0.2 y 0.5)

**Resultados:** El factor de aprendizaje de 0,7 alcanzo la estabilidad por un periodo de (3,6 s), lo que implica que es el equilibrio sostenido mas largo entre los identificados. Con relación al factor de descuento el mas destacable es el 0,2 , es el que establece el equilibrio mas optimo

## Artículo 5

**Título:** Reinforcement learning approach to control an inverted pendulum: A general framework for educational purposes

**Aporte:** Este artículo presenta un marco general para el control de un péndulo invertido utilizando algoritmos de aprendizaje por refuerzo (Reinforcement Learning, RL), destacándose por su aplicabilidad en entornos educativos. Proporciona una guía detallada para implementar los métodos Q-Learning y Deep Q-Networks (DQN) tanto en simulaciones como en experimentos reales. Además, aborda desafíos clave, como el ruido y la fricción, a través de simulaciones que replican las condiciones del mundo real, y valida la robustez de estos métodos para condiciones iniciales variadas. Al incluir un repositorio de código abierto, el trabajo fomenta la experimentación y el aprendizaje práctico de tecnologías avanzadas de control en un contexto académico, abriendo oportunidades para explorar el aprendizaje automático en aplicaciones dinámicas complejas.

**Proceso:** El estudio combina simulaciones y experimentos físicos para implementar el control del péndulo invertido mediante RL. Primero, se desarrolló un modelo físico del sistema y se definió un entorno RL que utiliza estados, acciones y recompensas para guiar el aprendizaje del agente. Se implementaron algoritmos Q-Learning y DQN, comparándolos en términos de eficiencia y precisión. Las simulaciones incluyeron variaciones en parámetros físicos clave como el voltaje, la fricción y el ruido, para evaluar el impacto en el aprendizaje y la estabilidad del control. Finalmente, se realizaron experimentos reales utilizando un sistema equipado con sensores y un motor controlado por un microcontrolador, validando los resultados obtenidos en simulaciones y destacando las ventajas de DQN frente a Q-Learning.

**Resultados:** Los resultados muestran que el algoritmo DQN logró estabilizar el péndulo invertido de manera eficiente y robusta frente a diversas condiciones iniciales y parámetros físicos, mientras que Q-Learning presentó limitaciones significativas en cuanto a tiempo de aprendizaje y precisión. En los experimentos, se identificó que un voltaje mínimo de 5.9 V es esencial para alcanzar un control óptimo, y las simulaciones confirmaron la efectividad del enfoque para manejar ruido y fricción moderados. Este trabajo destaca la capacidad de DQN para generalizar y adaptarse a condiciones variables, ofreciendo un marco educativo sólido para explorar aplicaciones de aprendizaje por refuerzo en sistemas dinámicos.

Capítulo 3: Metodología

* 1. Metodología de la Investigación
     1. Planificación de la revisión
     2. Desarrollo de la revisión
     3. Resultados de la revisión

Capítulo 4: Desarrollo

* 1. Desarrollo del aporte

Capítulo 5: Validación y resultados

Capítulo 6: Conclusiones y trabajos futuros

Referencias

Apellidos, Nombre (Año). Título del artículo. *Título del diario*, páginas desde - hasta.

Apellidos, Nombre (Año). *Título del libro.* Nombre de la ciudad: Nombre del editor

Notas al pie

1Agregue notas al pie, si corresponde, en su propia página después de las referencias. En el cuerpo de una nota al pie, como en este ejemplo, se usa el estilo de texto Normal. (Nota: Si elimina esta nota al pie de ejemplo, no se olvide de eliminar también su referencia en el texto. Está al final del párrafo Título 2 de ejemplo de la primera página del contenido del cuerpo de esta plantilla.)

Tablas

Tabla 1

Título de tabla

| Encabezado de columna | Encabezado de columna | Encabezado de columna | Encabezado de columna | Encabezado de columna |
| --- | --- | --- | --- | --- |
| Encabezado de fila | 123 | 123 | 123 | 123 |
| Encabezado de fila | 456 | 456 | 456 | 456 |
| Encabezado de fila | 789 | 789 | 789 | 789 |
| Encabezado de fila | 123 | 123 | 123 | 123 |
| Encabezado de fila | 456 | 456 | 456 | 456 |
| Encabezado de fila | 789 | 789 | 789 | 789 |

Nota: Coloque todas las tablas del documento en una sección de tablas, después de las referencias (y, si corresponde, después de las notas al pie). Use una página nueva para cada tabla e incluya un número de tabla y un título de tabla para cada una, como se muestra en esta página. Todo el texto explicativo aparece en una nota de tabla después de la tabla, como en esta. Use el estilo de tabla o ilustración, disponible en la galería de estilos de la pestaña Inicio, para agregar el espaciado entre la tabla y la nota. Las tablas en el formato de APA pueden usar un interlineado de una línea o de 1,5 líneas. Incluya un título para cada fila o columna, incluso si el contenido parece obvio. Se configuró un estilo de tabla para esta plantilla que cumple con las normas del estilo APA. Para insertar una tabla, seleccione la pestaña Insertar y haga clic en Tabla.

Ilustraciones



Ilustración 1. Incluya todas las ilustraciones en su propia sección, después de las referencias (y, si corresponde, después de las notas al pie y las tablas). Incluya un título numerado para cada ilustración. Use el estilo de tabla o ilustración para agregar fácilmente espaciado entre la ilustración y el título.

Para obtener más información sobre todos los elementos del formato de estilo APA, vea el Manual de estilo de la APA, 6.ª edición.