

Título: Diplomado en Machine Learning y Procesamiento del Lenguaje Natural			
Duración	60 horas	Modalidad	Virtual Sincrónica
Objetivo General: Fortalecer las competencias técnicas y analíticas del equipo de tecnología del Centro Naval mediante el estudio de los fundamentos y aplicaciones prácticas del Machine Learning (ML) y el Procesamiento del Lenguaje Natural (PLN), con énfasis en la seguridad, defensa y análisis de información estratégica.			

2. Estructura Curricular del Diplomado (Módulos)

Módulo	Nombre del Módulo	Duración (Horas)
1	Fundamentos de Inteligencia Artificial y Machine Learning	6
2	Preparación y análisis de datos	12
3	Modelado con aprendizaje supervisado	15
4	Modelado con Aprendizaje no supervisado	3
5	Introducción al procesamiento de lenguaje natural (PLN)	6
6	Introducción a la operacionalización y despliegue de modelos	3
7	Proyecto final integrador	15
Total		60

En un entorno cada vez más dinámico, digitalizado y basado en datos, el aprovechamiento estratégico de tecnologías como la inteligencia artificial se convierte en una necesidad operativa para instituciones clave en la defensa y seguridad nacional. La Armada Nacional, en su rol de protección de la soberanía marítima y de gestión de información crítica, demanda soluciones cada vez más inteligentes, predictivas y eficientes para la toma de decisiones informadas en tiempo real.

Este proceso formativo está diseñado para que los participantes tengan una visión general de cómo resolver problemas con IA, ML o PLN, desde la lógica del ciclo de vida de un proyecto de ciencia de datos. Se iniciará con los fundamentos metodológicos y

la preparación de datos, para luego avanzar hacia el entrenamiento y la evaluación de modelos de aprendizaje supervisado y no supervisado. Posteriormente, se introducirán áreas especializadas como el PLN, culminando con la capacidad de operacionalizar un modelo a través de una API (Interfaz de Programación de Aplicaciones). El diplomado prioriza el uso de bibliotecas modernas y entornos de desarrollo abiertos, potenciando así la autonomía y escalabilidad de los proyectos generados.

Más allá de la transferencia de conocimiento técnico, este proceso formativo busca fomentar una mentalidad analítica y de innovación orientada a la resolución de problemas. El objetivo final es preparar al equipo para que pueda no solo entender, sino también construir y desplegar prototipos funcionales, liderando así transformaciones tecnológicas alineadas con la misión institucional de la Armada y con los estándares globales en ciencia de datos aplicada a la defensa.

3. Detalle de Módulos

Módulo 1: Fundamentos de Inteligencia Artificial y Machine Learning (6 horas)

- **Sesión 1:** Introducción a la IA: historia, evolución y aspectos éticos en entornos militares.
- **Sesión 2:** Ciclo de vida de un proyecto de ML y el equipo de ciencia de datos.

Resultado esperado: Comprensión del marco conceptual de la IA, su ciclo de vida, y el rol estratégico que juega en el ámbito militar.

Módulo 2: Preparación y análisis de datos (12 horas)

- **Sesión 3:** Preparación e instalación de las herramientas necesarias para conformar un ecosistema de desarrollo, enfocado a resolver problemas de ciencia da datos.
- **Sesión 4:** Análisis exploratorio de datos (EDA)
- **Sesión 5:** Técnicas de limpieza y preprocesamiento de datos.
- **Sesión 6:** Ingeniería y selección de características (*feature engineering*).

Resultado esperado: Competencias para explorar, limpiar y transformar datos crudos en un conjunto de datos estructurado y optimizado para el modelado.

Módulo 3: Modelado con aprendizaje supervisado (15 horas)

- **Sesión 7:** Modelos de regresión I: enfoque basado en modelos lineales.
- **Sesión 8:** Modelos de regresión II: enfoque basado en árboles de decisión.
- **Sesión 9:** Modelos de clasificación I: enfoque basado en modelos lineales.
- **Sesión 10:** Modelos de clasificación II: enfoque basado en árboles de decisión.
- **Sesión 11:** Evaluación y optimización de modelos: validación cruzada y afinamiento de hiperparámetros.

Resultado esperado: Capacidad para implementar, evaluar y comparar diferentes algoritmos de aprendizaje supervisado para resolver problemas de regresión y clasificación.

Módulo 4: Aprendizaje no supervisado (3 horas)

- **Sesión 12:** Segmentación de datos con el algoritmo *k-means*.

Resultado esperado: Habilidad para aplicar técnicas de clustering para la segmentación de datos y la extracción de insights a partir de grupos no etiquetados.

Módulo 5: Introducción al procesamiento de lenguaje natural (PLN) (6 horas)

- **Sesión 13:** Fundamentos de PLN y preprocesamiento de texto.
- **Sesión 14:** Construcción de un clasificador de texto end-to-end con TF-IDF (*Term Frequency-Inverse Document Frequency*).

Resultado esperado: Capacidad para construir una solución básica de clasificación de texto, comprendiendo el flujo de trabajo específico del procesamiento de lenguaje natural.

Módulo 6: Operacionalización de modelos (3 horas)

- **Sesión 15:** Despliegue básico de modelos como API local.

Resultado esperado: Competencia para encapsular un modelo de ML entrenado en una API local, dando el primer paso hacia la operacionalización.

Módulo 7: Proyecto final integrador (15 horas)

Este módulo se desarrolla bajo una metodología de aprendizaje basado en proyectos (ABP). Las sesiones son talleres prácticos donde el instructor actúa

como mentor, guiando a los participantes en el desarrollo de sus soluciones. La evaluación se centra en el prototipo funcional y la presentación final.

- **Sesión 16: Lanzamiento del proyecto y planificación.** Definición del alcance final de los proyectos, formación de equipos (si aplica), y estructuración de un plan de trabajo (sprint plan). Se establecen los criterios de éxito y los entregables.
- **Sesión 17: Taller de desarrollo I - Construcción del prototipo.** Sesión de trabajo práctico enfocado en integrar los modelos de ML/PLN desarrollados en una solución cohesiva. El instructor ofrece soporte técnico y conceptual.
- **Sesión 18: Taller de desarrollo II - Refinamiento y storytelling con datos.** Los participantes finalizan sus prototipos. Se introduce el concepto de "storytelling" para preparar una presentación que comunique eficazmente el problema, la solución y el impacto a una audiencia técnica y no técnica.
- **Sesiones 19 y 20: Presentación de proyectos finales y cierre.** Exposición de los proyectos. Cada presentación debe incluir una demostración del prototipo, un análisis de los resultados y una discusión sobre su aplicabilidad y valor estratégico.

Resultado esperado: Desarrollo de un prototipo funcional de ML/PLN que soluciona un problema definido, y la habilidad para comunicar sus resultados y su impacto potencial.

4. Metodología

- Enfoque práctico basado en casos de uso reales del entorno militar
- Desarrollo en entornos de simulación y datos sintéticos
- Acompañamiento permanente del formador
- Evaluación por entregas parciales y presentación final

5. Resultados Esperados

- Fortalecimiento de capacidades analíticas y predictivas
- Habilidades en la creación de soluciones con datos estructurados y no estructurados
- Apropiación de herramientas de IA en el contexto naval