Fundamentos algorítmicos y ética en IA: de la teoría a la práctica

Kit de herramientas para sesiones sincrónicas

Ejercicios basados en proyectos

Número de proyecto| 2022-1-ES01-KA220-HED-000085257

Ejercicios basados en proyectos

Ejercicios basados en proyectos

	Descripción
Descripción de la tarea	En equipos de 4 o 5 personas, trabajarán en la evaluación de una de las seis soluciones de IA en el transcurso de cinco unidades del curso. Esta tarea integral implica un análisis detallado de la solución de IA desde múltiples perspectivas, siguiendo la estructura del curso y los temas tratados en cada unidad del curso.
Descripción de cómo realizar la tarea	**Instrucciones** • **Su equipo trabajará en una solución de IA de la lista de la página 4.** **Analice la solución elegida en función de las cinco perspectivas proporcionadas en la página 6** **Documente sus hallazgos en un informe estructurado, abordando cada criterio de evaluación en detalle. Cada punto debe presentarse en un máximo de 2 diapositivas** **Presente su análisis a la clase en la última sesión sincrónica, destacando las ideas clave y las recomendaciones para mejorar la solución.** **Su presentación al equipo debe durar entre 5 y 10 minutos**
Tiempo estimado para realizar la tarea	Approx. 1-2 hrs / CU (5-10 hrs total)
Sugerencia de fuentes para realizar la tarea	Manual interactivo y otro material del curso + cualquier otra fuente
Descripción detallada de cómo entregar la tarea	Ver páginas siguientes.
Información sobre la fecha límite para la entrega de la tarea	Presentado en la última sesión sincrónica.
Información de contacto o cómo aclarar dudas	El profesor debe proporcionar una forma de contacto (puede ser una dirección de correo electrónico, un número de teléfono...).

Ejercicios basados en proyectos
Cronología e instrucciones

CRONOGRAMA DEL EJERCICIO DEL PROYECTO

Presentación del proyecto y formación de equipos.

- Preferiblemente, comenzar a trabajar en el proyecto después de la 1ª sesión sincrónica, acordar con el equipo reuniones para trabajar en el proyecto. Alternativamente, comience después de la 2ª sesión síncrona que cubre las CU 1 y 2.

Revisar el progreso

En la 3ª y 4ª sesiones sincrónicas, posibilidad de revisar el progreso y discutir cualquier desafío con el profesor..

Presentación Final

En la 5ª sesión sincrónica presentar el ejercicio del proyecto en equipo.
Presentación 5-10 minutos.
Puedes elegir el diseño de PowerPoint o Canva, recuerda hacer una presentación atractiva.

Más instrucciones

- El tema se presenta en la siguiente diapositiva. Como equipo, podrán decidir el nombre, el propósito y el contenido más detallado de su servicio. Sin embargo, la solicitud debe estar en el campo que se le proporciona en la página siguiente.
  Continúe trabajando en el proyecto al ritmo acordado con el equipo del proyecto, sin embargo, el contenido siempre está relacionado con las unidades del curso, por lo que es preferible trabajar después de que los miembros del grupo hayan finalizado los aprendizajes electrónicos por tema.
  Puedes decidir cómo construir tu presentación; Por lo tanto, las preguntas están ahí como una guía, pero debe crear un contenido que sea atractivo para la audiencia.

NOTA | En caso de que el estudiante no esté disponible para el trabajo en grupo (o no participe activamente en el proyecto), el ejercicio del proyecto debe entregarse individualmente al profesor (10 diapositivas)

Ejercicios basados en proyectos
Soluciones de IA divididas por equipo

La solución de IA en la que debería trabajar tu equipo:

Herramienta de diagnóstico sanitario impulsada por IA
A tool that assists doctors by providing preliminary diagnoses based on patient symptoms and medical history.
- XX grupos que trabajan sobre el tema (añadir aquí los nombres de los alumnos, divididos en grupos de 4-5 por el profesor)

Sistema de recomendación de servicios de streaming
Una aplicación que sugiere películas y series a los usuarios en función de su historial de visualización y preferencias.
- XX grupos que trabajan sobre el tema (añadir aquí los nombres de los alumnos, divididos en grupos de 4-5 por el profesor)

Sistema domótico inteligente
- Una solución que automatiza los dispositivos domésticos, como luces, termostatos y sistemas de seguridad, en función de las rutinas y preferencias del usuario.
  - XX grupos que trabajan sobre el tema (añadir aquí los nombres de los alumnos, divididos en grupos de 4-5 por el profesor)

Asesor de inversiones financieras impulsado por IA
- Un asesor que proporciona recomendaciones de inversión basadas en datos de mercado y objetivos financieros del usuario.
  - XX grupos que trabajan sobre el tema (añadir aquí los nombres de los alumnos, divididos en grupos de 4-5 por el profesor)

Aplicación de herramienta de compra en línea
- Una aplicación que recomienda productos a los usuarios en función de su historial de navegación y compras.
  - XX grupos que trabajan sobre el tema (añadir aquí los nombres de los alumnos, divididos en grupos de 4-5 por el profesor)

Plataforma de e-Learning personalizada
Una plataforma que personaliza las rutas de aprendizaje y el contenido para los estudiantes en función de sus estilos de aprendizaje, progreso y preferencias.
- XX grupos que trabajan sobre el tema (añadir aquí los nombres de los alumnos, divididos en grupos de 4-5 por el profesor)

Ejercicios basados en proyectos
UC 1 CONTENIDO RELACIONADO

Comience definiendo lo siguiente(Esta tarea no debe ocupar más de 2 diapositivas de presentación)

Planteamiento del problema: ¿Qué problema está resolviendo su aplicación?
Objetivo: ¿Cuál es el objetivo de su aplicación de IA? ¿Cómo se llama la aplicación? ¿Cuáles son los criterios de éxito de su solución?

Crear un mapa de partes interesadas, que debe definir al menos las partes interesadas en el uso, las partes interesadas externas y las partes interesadas en el desarrollo (utiliza el material de la UC1 como guía).

Considere qué tan bien se cumplen los siguientes principios éticos con su solicitud, el enfoque principal debe estar en los aspectos sociales y ambientales:

Beneficencia:
- Garantizar que la solución de IA mejore el bienestar individual y el bien social. Áreas clave a tener en cuenta:
  Protección de los derechos fundamentales: salvaguardar los derechos individuales, reforzando la integridad ética.
  Sostenibilidad ambiental: minimizar el impacto ambiental a través de prácticas sostenibles.
  Justificación: definir y vincular claramente el propósito del sistema con los beneficios tangibles de la sociedad.

NEn-maleficencia:
- Asegurar que la identificación del problema y la solución propuesta no perjudiquen a las personas o a las comunidades.
  Considere y monitoree cuidadosamente el impacto del sistema en el bienestar físico y mental.
  Evaluar los posibles riesgos de seguridad y las estrategias de mitigación.

Justicia:
Evaluar el impacto del sistema en las instituciones, la democracia y la sociedad.

Ejercicios basados en proyectos

UC 1 (Diapositiva de ayuda)

Utilice esta tabla para ver más en detalle las posibles consideraciones bajo cada principio ético, combinado por Morley *et. àl* (2020)

Table 5 showing the connection between high-level ethical principles and tangible system requirements as adapted from the methodology outlined in Chapter II of the European Commission's "Ethics Guidelines for Trustworthy AI"

Principle	Beneficence	Non-Maleficence	Autonomy	Justice	Explicability
Requirements	**Stakeholder participation:** to develop systems that are trustworthy and support human flourishing, those who will be affected by the system should be consulted **Protection of fundamental rights** **Sustainable and environmentally friendly AI:** the system's supply chain should be assessed for resource usage and energy consumption **Justification:** the purpose for building the system must be clear and linked to a clear benefit—system's should not be built for the sake of it	**Resilience to attack and security:** AI systems should be protected against vulnerabilities that can allow them to be exploited by adversaries. **Fallback plan and general safety:** AI systems should have safeguards that enable a fallback plan in case of problems. **Accuracy:** for example, documentation that demonstrates evaluation of whether the system is properly classifying results. **Privacy and Data Protection:** AI systems should guarantee privacy and data protection throughout a system's entire lifecycle. **Reliability and Reproducibility:** does the system work the same way in a variety of different scenarios. **Quality and integrity of the data:** when data is gathered it may contain socially constructed biases, inaccuracies, errors and mistakes—this needs to be addressed **Social Impact:** the effects of system's on people's physical and mental wellbeing should be carefully considered and monitored	**Human agency:** users should be able to make informed autonomous decisions regarding AI systems **Human oversight:** may be achieved through governance mechanisms such as human-on-the-loop, human-in-the-loop, human-in-command	**Avoidance of unfair bias** **Accessibility and universal design** **Society and democracy:** the impact of the system on institutions, democracy and society at large should be considered **Auditability:** the enablement of the assessment of algorithms, data and design processes. **Minimisation and reporting of negative impacts:** measures should be taken to identify, assess, document, minimise and respond to potential negative impacts of AI systems **Trade-offs:** when trade-offs between requirements are necessary, a process should be put in place to explicitly acknowledge the trade-off, and evaluate it transparently **Redress:** mechanism should be in place to respond when things go wrong	**Traceability:** the data sets and the processes that yield the AI system's decision should be documented **Explainability:** the ability to explain both the technical processes of an AI system and the related human decisions **Interpretability**

FUENTE| Morley, J., Floridi, L., Kinsey, L., & Elhalal, A. (2020). From What to How: An Initial Review of Publicly Available AI Ethics Tools, Methods and Research to Translate Principles into Practices. Science and Engineering Ethics, 26(4), 2141–2168. https://doi.org/10.1007/s11948-019-00165-5

Ejercicios basados en proyectos
CONTENIDO RELACIONADO de las UCs

UC 2
Preocupaciones sobre la privacidad

- Identifique los problemas de privacidad asociados con la solución.

Tenga en cuenta cómo se recopilan, almacenan y utilizan los datos.
Sugerir medidas para proteger la privacidad del usuario y garantizar la seguridad de los datos.

UC 3
Algoritmo utilizado

- Describa el algoritmo que impulsa la solución.
  Explique cómo se construye y entrena el algoritmo en función de los datos recopilados (cu2).
  Discuta cualquier detalle técnico específico que sea importante para comprender la solución.

UC 4
Posibles sesgos

- Identificar posibles sesgos en el sistema.

Analice qué podría causar estos sesgos (p. ej., datos de entrenamiento, diseño de algoritmos).
Sugerir formas de prevenir o mitigar estos sesgos.

UC 5

- La Universidad Técnica de Delft ha creado una lista de comprobación para el desarrollo ético de la IA. Verifique la herramienta a través de este link y utilice la lista de verificación del proyecto AI APP en las páginas 12-13 y complete la lista de verificación en la página 13. Presente su lista de verificación al final de la presentación de su proyecto.

- La diapositiva final debe presentar lo que ustedes como equipo aprendieron y qué nuevos conocimientos obtuvieron al hacer esta presentación.

GRACIAS

Número de proyecto: 2022-1-ES01-KA220-HED-000085257

Cofinanciado por la Unión Europea

El apoyo de la Comisión Europea a la producción de esta publicación no se compone de los contenidos, que reflejan únicamente las opiniones de los autores, y la Comisión no se hace responsable del uso que pueda hacerse de la información contenida en ella.