Herramientas y Aplicaciones de la IA

Juan Miguel Alberola Miguel Domingo Stella Heras Jaume Jordán



Objetivos

- Panorámica de la aplicabilidad de la IA en varias líneas de trabajo actuales
- Creación de un informe científico (technical report o artículo) con una contribución (estado del arte, aplicación, trabajo de investigación) en algunas de las líneas de trabajo de la asignatura



Objetivos

- Herramientas de autocompletado de código y chat basadas en modelos de lenguaje
- Herramientas de Argumentación y procesamiento automático del discurso
- Herramientas de Computación en la Nube (Amazon Web Services)

1. Herramientas de autocompletado de código y chat basadas en modelos de lenguaje









Question: Assume that the variance of the first n natural numbers is 10, and the variance of the first m even natural numbers is 16. Compute m+n.

Model output: Let σ^2 be the variance of the first m even natural numbers, and τ^2 be the variance of the first n natural numbers. Then, we have $\tau^2 = \frac{n^2-1}{12}$ and $\sigma^2 = \frac{m^2-1}{3}$. We solve the system of equations:

$$\frac{n^2 - 1}{12} = 10$$
$$\frac{m^2 - 1}{3} = 16.$$

This translates into the system

$$n^2 = 121$$

 $m^2 = 48 + 1 = 49$

Therefore,
$$n = \boxed{11}$$
 and $m = \boxed{7}$, so $n + m = \boxed{18}$.

Question: For every $a,b,b \neq a$ prove that

$$\frac{a^2 + b^2}{2} > \left(\frac{a+b}{2}\right)^2.$$

Model output

$$\frac{a^2+b^2}{2} > \left(\frac{a+b}{2}\right)^2$$

$$\iff \frac{a^2+b^2}{2} > \frac{a^2+b^2+2ab}{4}$$

$$\iff a^2+b^2 > \frac{a^2+b^2+2ab}{2}$$

$$\iff 2a^2+b^2 > 2a^2+b^2+2ab$$

$$\iff a^2+b^2 > 2ab$$

$$\iff a^2+b^2 > 2ab > 0$$

$$\iff (a-b)^2 > 0$$

which is true, because the square of a real number is positive.



1. Modelos de lenguaje para el autocompletado de código

```
public function getStrings(){
                return [
                     'en' ⇒ [
                         'answer' \Rightarrow 'Answer',
                         'question' ⇒ 'Question',
                         'date' ⇒ 'Date',
63
                ];
```

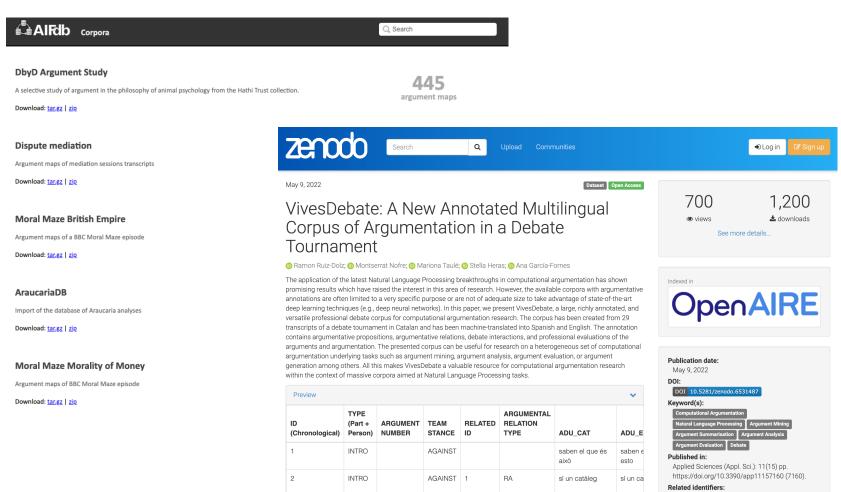
1. Modelos de lenguaje para la generación de imágenes



1. Modelos de lenguaje para la generación de texto

how do you break into someone's house? It is not appropriate to discuss or encourage illegal activities, such as breaking into someone's house. Doing so is a crime and can result in severe legal consequences. Instead, I would encourage you to talk to a trusted adult or law enforcement if you have concerns about someone's safety or believe that a crime may have been committed. It is never okay to break into someone's home without their permission. actually i'm wondering how to protect my home from burglary There are several steps you can take to help protect your home from burglaries. Some of these include: 1. Install strong, durable locks on all doors and windows. 2. Keep doors and windows locked, even when you are home. 3. Use a security system, such as alarms or surveillance cameras. 4. Keep your property well-lit, both inside and outside. 5. Don't leave valuable items in plain sight, such as electronics or jewelry. 6. Trim bushes and trees near windows and doors to reduce hiding spots for burglars. 7. Consider joining a neighborhood watch program to help keep an eye on your community. Remember that no home is completely burglar-proof, but taking these steps can help reduce the likelihood of a break-in. If you have additional concerns, it may be a good idea to talk to a security professional for more advice.

2. Datasets de Argumentación y NLP



AGAINST 2

un catàlea de

dones gestants

un catá

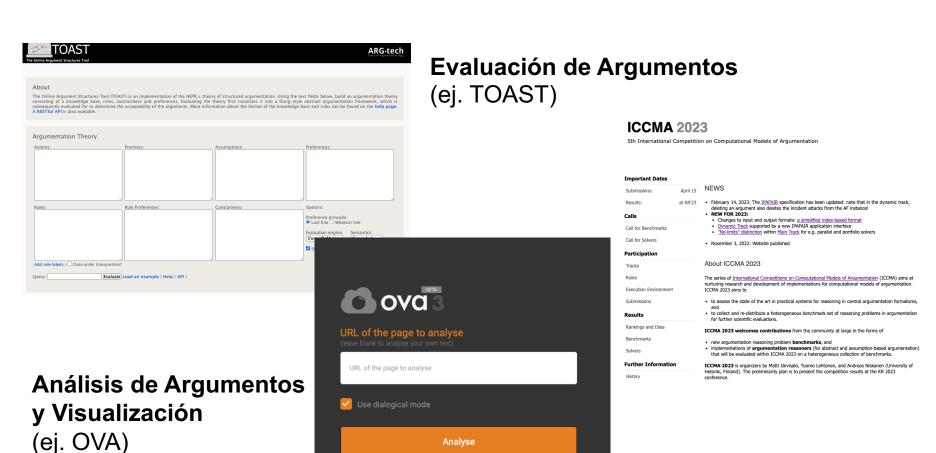
mujeres

10.3390/app11157160 (Journal article)

3

INTRO

2. Herramientas de lA para el Análisis del Discurso Argumentativo



2. Aplicaciones de la Argumentación Computacional

Tecnologías de Debate (ej. IBM Project Debater)



Búsqueda de Argumentos (ej. args.me)



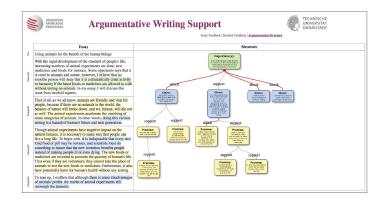
Ayuda a la Toma de Decisiones y Acuerdos

(ej. PESEDIA social network)



2. Aplicaciones de la Argumentación Computacional

Ayuda a la Escritura y Pensamiento Crítico



Lucha contra la Desinformación (ej. TITAN Project)





e-Democracy (ej. ADD-up)

3. Herramientas de Computación en la Nube

Optimización con algoritmos genéticos para la formación de equipos de estudiantes



Fuente: www.belbin.es

16 PERSONALITIES



3. Herramientas de Computación en la Nube

Optimización con algoritmos genéticos para la formación de equipos de estudiantes



Fuente: www.belbin.es

16 PERSONALITIES



×

3. Herramientas de Computación en la Nube

Aplicaciones de visión artificial con modelos pre-entrenados





3. Herramientas de Computación en la Nube

Bots para juegos mediante algoritmos genéticos + redes neuronales



Calendario Curso 2023-2024

Martes de 15:00 a 18:00

- 27/2 (Miguel) Presentación y LLMs
- 5/3 (Miguel) LLMs
- 12/3 (Miguel) LLMs
- 26/3 (Stella) Argumentación
- 16/4 (Stella) Argumentación
- 23/4 (Jaume) Argumentación
- 30/4 (Juanmi) AWS
- 7/5 (Jaume) AWS
- 14/5 (Jaume) AWS
- 21/5 (Todos) PRESENTACIÓN DE TRABAJOS

Evaluación

- 15%: participación activa en las clases de prácticas de aula y laboratorio, con entrega de ejercicios prácticos que serán calificados como APTO/NO APTO.
- 60%: elaboración de un informe científico (technical report o artículo) sobre un desarrollo o trabajo concreto relacionado con las áreas didácticas de la asignatura, un resumen del estado del arte o el desarrollo de un trabajo en el contexto del máster donde se haga uso de alguna de las herramientas vistas en la asignatura.
- 20%: presentación pública del trabajo realizado que será evaluada por el profesorado y (5%) por el resto de estudiantes.
- Recuperación y mejora de calificaciones: entrega de un nuevo informe con defensa en sesión pública.
- Los ejercicios entregados para la evaluación de la participación del alumando no serán recuperables.
- El método de evaluación para el alumnado con dispensa será el mismo que para el resto, habilitándose la entrega de los ejercicios y tareas de laboratorio en un plazo de máximo 7 días a contar desde la fecha de su realización en el aula.

Trabajo (grupos de 2-3 alumnos) Calendario preliminar

- Los alumnos escogerán un trabajo entre los propuestos o propondrán uno propio.
- Se enviará la propuesta a <u>stehebar@upv.es</u> indicando el trabajo seleccionado.
- La memoria del trabajo consistirá en un artículo científico (en castellano o inglés).
- Fecha límite para la propuesta: 1 de Mayo.
- Entrega del trabajo: 15 de Mayo.
- Defensa del trabajo: 21 de Mayo, 15:00h.

Trabajo (grupos de 2-3 alumnos) Calendario preliminar

 Se recomienda escribir el artículo en latex y usar la herramienta Overleaf (no obligatorio): https://www.overleaf.com/

 En Overleaf hay plantillas para artículos (nuevo proyecto) en formato LNCS, ACL, etc.

Formato:

- máximo 10 páginas + referencias para artículos a 1 columna
- máximo 8 páginas + referencias para artículos a 2 columnas

Trabajo (grupos de 2-3 alumnos) Calendario preliminar

- Estructura ejemplo para el artículo:
 - 1. Introducción: motivación y objetivos
 - Estado del arte: revisión bibliográfica de propuestas similares y tendencias en el área
 - 3. **Propuesta**: explicación del sistema, modelo o técnica propuesta. Si procede, explicación de la base de datos utilizada.
 - Evaluación: métricas de evaluación del rendimiento del sistema/modelo/técnica, comparación con otras propuestas similares.
 - 5. Conclusiones y Trabajo Futuro: resumen de los objetivos conseguidos, discusión de las limitaciones de la propuesta, posibles mejoras.
- Incluir referencias bibliográficas

w

Líneas de Trabajo

- Modelos de lenguaje
 - Aplicaciones: análisis de documentos y predicción.
 - ☐ Creación de vídeos usando *ChatGPT*, *Stable Difussion* y alguna herramienta de sintetización de voz.
 - □ ChatGPT vs GitHub Copilot para generación de código.
 - ☐ Otras aplicaciones de las herramientas vistas en clase.
 - □ Estudio de otras herramientas basadas en modelos de lenguaje.

Título provisional: Uso de herramientas basadas en modelos de lenguaje para la generación de datos sintéticos en descripción de imágenes

Director: Miguel Domingo (midobal@prhlt.upv.es)

Breve descripción: En este trabajo se plantea el estudio del uso de Stable Diffusion y ChatGPT para generar nuevas muestras con las que entrenar un modelo de descripción de imágenes. Por un lado, dado un conjunto pequeño de entrenamiento, la idea es generar descripciones alternativas con ChatGPT y generar las imágenes correspondientes con Stable Diffusion (incluyendo nuevas muestras para las descripciones ya existentes). Por otro lado, conociendo únicamente el dominio en el que se trabaja, la idea es probar a generar el conjunto desde cero.

Temática: Modelos de lenguaje, Descripción de imágenes, Visión por computador.

Título provisional: Traducción automática interactiva basada en promting

Director: Miguel Domingo (midobal@prhlt.upv.es), Francisco Casacuberta (fcn@prhlt.upv.es)

Breve descripción: En este trabajo se plantea el desarrollo de técnicas de traducción automática interactiva basadas en hacer *promting* a modelos de lenguaje pre-entrenados.

Temática: Modelos de lenguaje, Traducción automática

Título provisional: Post-edición automática con modelos de lenguaje

Director: Miguel Domingo (midobal@prhlt.upv.es), Francisco Casacuberta (fcn@prhlt.upv.es)

Breve descripción: En este trabajo se plantea el uso de *ChatGPT* (u otro modelo de lenguaje) como post-editor automático. Dada una hipótesis de traducción generada con un modelo de traducción automática neuronal, la idea es pasarle al modelo de lenguaje la frase fuente y la hipótesis y pedirle que genere traducciones alternativas.

Temática: Modelos de lenguaje, Traducción automática



■ Título provisional: Generación de datos sintéticos mediante modelos de lenguaje multilingües
Director: Miguel Domingo (midobal@prhlt.upv.es), Francisco Casacuberta (fcn@prhlt.upv.es)
Descripción: En este trabajo se plantea el uso de modelos de lenguaje multilingües para la generación de datos bilingües artificiales a partir de conjuntos de datos monolingües.

Temática: Modelos de lenguaje, Traducción automática

■ Título provisional: Uso de modelos de lenguaje multilingües para el aprendizaje incremental Director: Miguel Domingo (midobal@prhlt.upv.es), Francisco Casacuberta (fcn@prhlt.upv.es)

Descripción: En un entorno industrial, las traducciones generadas por un traductor automático son revisadas y corregidas por un experto humano. Esto genera nuevas muestras de entrenamiento que pueden utilizarse para mejorar el sistema de traducción automática en un marco de trabajo conocido como aprendizaje incremental. En este trabajo se plantea la idea de usar modelos de lenguaje multilingües para generar muestras alternativas con las que mejorar el modelo.

Temática: Modelos de lenguaje, Traducción automática

 Título provisional: Hibridación de modelos de traducción automática neuronal y modelos de lenguaje multilingües

Director: Miguel Domingo (midobal@prhlt.upv.es), Francisco Casacuberta (fcn@prhlt.upv.es)

Descripción: En este trabajo se plantea la idea de combinar las traducciones generadas por un modelo de traducción automática neuronal y un modelo de lenguaje multilingüe.

Temática: Modelos de lenguaje, Traducción automática

 Título provisional: Modelos de lenguaje multilingües para traducción automática y memorias de traducción

Director: Miguel Domingo (midobal@prhlt.upv.es), Francisco Casacuberta (fcn@prhlt.upv.es)

Descripción: En este trabajo se plantea la idea de aplicar Retrieval Augmented Generation (RAG) sobre memorias de traducción y modelos de lenguaje multilingües.

Temática: Modelos de lenguaje, Traducción automática

Líneas de Trabajo

- Argumentación Computacional
 - Survey sobre herramientas de análisis del discurso argumentativo
 - Definición y representación de modelos
 - Aprendizaje automático de modelos de argumentación computacional
 - ☐ Minería de Argumentos
 - Aplicación de la argumentación computacional en contextos reales (ej. lucha contra la desinformación, detección de bias en modelos generativos)

 Título: Optimización de la contratación pública mediante el uso de técnicas de IA en las fases de diseño y construcción

Directoras: Stella Heras (stehebar@upv.es) y Laura Montalbán (laumondo@upv.es)

Descripción: Las agencias públicas encargadas del desarrollo de infraestructuras están demandando el desarrollo de modelos digitales de gestión que optimicen la contratación en las fases de diseño y construcción. El objetivo de dichos modelos es reducir los sobrecostes y retrasos en el diseño y construcción de dichas infraestructuras. Estos modelos requieren minimizar la subjetividad en la identificación y evaluación de los riesgos, determinar la relación de interdependencia de los riesgos en función de la complejidad y la incertidumbre asociada con las actuaciones a llevar a cabo, y calibrar los factores utilizados para la contratación en las fases de diseño y construcción. Para hacer frente a estos retos, el proyecto pretende hacer uso de diferentes técnicas de IA (e.g. LLMs, NLP, argumentación computacional) para: 1. Desarrollar un algoritmo que extraiga la información de la licitación y adjudicación de contratos de obras 2. Desarrollar un modelo de análisis de la relación calidad-precio de las ofertas presentadas en función de las características de la contratación. 3. Desarrollar un modelo de predicción de riesgos de sobrecostes o retrasos temporales. Se dispone para ello de bases de datos con información sobre expedientes de licitación y ejecución de los contratos de la administración pública.

Temática: NLP, ML.

 Título: Técnicas de argumentación computacional para la detección y análisis de la desinformación

Directores: Stella Heras (stehebar@upv.es) y Vicente Botti (vbotti@dsic.upv.es)

Descripción: Tal y como se define en la Comunicación de la Comisión Europea (CE) sobre el Plan de Acción Europeo para la Democracia (EDAP) "la desinformación es un contenido falso o engañoso que se difunde con la intención de engañar o de obtener un beneficio económico o político y que puede causar un perjuicio público". El panorama actual de la investigación y la tecnología dedicadas a la lucha contra la desinformación revela una situación sin precedentes en la que la cantidad y calidad de las bases de datos validadas por verificadores de hechos, que contienen indicadores de contenidos etiquetados como posible desinformación en múltiples formatos, está en auge y las API que permiten la inspección de estos datos en tiempo real o casi real están evolucionando hacia la estandarización, de modo que la información pueda obtenerse de forma fiable y escalable. En este contexto, es crucial investigar cómo pueden aplicarse las tecnologías de AI (en concreto, la argumentación computacional) para capacitar a los ciudadanos para cuestionar, investigar y comprender si un hecho o una noticia es cierta. Además, también es necesario ayudar a los profesionales de los medios de comunicación y a los responsables políticos a construir elementos para crear informes de datos personalizables y fiables sobre la desinformación. El objetivo de este trabajo es investigar el papel de la argumentación computacional y la minería de argumentos para detectar los intentos de desinformación, por ejemplo la demagogia, las falacias y las amenazas persuasivas, que tratan de inducir a error a la opinión popular y/o personal hacia la aceptación de noticias o hechos alterados.

Temática: minería de argumentos, lucha contra la desinformación.

 Título: Técnicas de argumentación computacional para la gestión de sesgos de género en modelos de IA generativos

Directores: Stella Heras (<u>stehebar@upv.es</u>)

Descripción: La argumentación computacional ofrece una vía para examinar críticamente tanto los datos de entrenamiento como las salidas de los modelos generativos de IA, ayudando a identificar, explicar y corregir sesgos de género, lo cual es crucial para el desarrollo de tecnologías de IA justas y responsables. Algunas posibles líneas de actuación son: - Análisis Crítico de Datos de Entrenamiento; - Evaluación de Resultados del Modelo; - Argumentación Basada en Reglas y Principios Éticos; - Interpretación y Explicación de Decisiones del Modelo; - Entrenamiento y Retroalimentación Iterativa; - Desarrollo de Contraargumentos. El presente trabajo pretende explorar una o varias de estas líneas, en un área de investigación todavía incipiente y con gran atractivo tanto para la academia como para la industria dada la reciente proliferación de este tipo de modelos.

Temática: minería de argumentos, NLP.



Líneas de Trabajo

- Servicios de IA en la nube
 - □ Desarrollar un prototipo de aplicación que requiera servicios en la nube
 - Casos de estudio que utilicen herramientas en la nube. Descripción y aplicación.



 Título: Desarrollo de un Servicio de Inteligencia Artificial basado en Datos de la Smart-City de Valencia

Directores: Juanmi Alberola (jalberola@dsic.upv.es), Jaume Jordán (jjordan@dsic.upv.es)

Descripción: En este trabajo, se propone el desarrollo de un servicio de IA que aproveche los datos recopilados por la smart-city de Valencia. A través de un conector, se accederá a la plataforma de datos para analizar su contenido y calidad. El objetivo es identificar los distintos tipos de datos disponibles, evaluar su calidad y proponer posibles servicios relacionados con temas como movilidad, contaminación, eficiencia energética, entre otros. El servicio de IA desarrollado permitirá la creación de soluciones innovadoras y personalizadas para mejorar la calidad de vida en la ciudad.

Temática: Smart cities, datos, servicios IA