



La Falacia del Porcentaje y un Marco Neutrosófico de Valor Único para la Indeterminación en los Veredictos de Detectores de IA sobre Autoría Académica

The Percentage Fallacy and a Neutrosophic One-Value Framework for Indeterminacy in AI Authorship Detector Verdicts

Oscar Caicedo-Camposano 1, *, Mayra Jara-Sen San 1, César Caicedo-Camposano 2 and Ariana Cruz-Posligua 2

1 Universidad Técnica de Babahoyo, Ecuador

2 Universidad Regional Autónoma de Los Andes, Ecuador

* Correspondence: ocamposano@utb.edu.ec

Resumen: Los detectores de contenido de inteligencia artificial (IA) han ingresado en la evaluación académica universitaria como instrumentos de verificación de autoría, frecuentemente sin regulación formal ni consideración de sus tasas de error documentadas. Este artículo propone un marco neutrosófico de valor único (SVN) para formalizar la indeterminación epistemológica que caracteriza los veredictos de estos sistemas. Argumentamos que todo detector de IA colapsa una tripleta neutrosófica implícita $\langle T, I, F \rangle$ en un escalar unidimensional, eliminando el componente de indeterminación I y produciendo lo que denominamos la falacia del porcentaje. Para resolver esta indeterminación, proponemos una función de puntuación neutrosófica $S = T + \alpha \cdot D - I - F$ que incorpora la maestría cognitiva demostrada D mediante defensa oral o evaluación directa, y formalizamos el Principio de Primacía Cognitiva (PPC) como criterio prevalente de autoría académica. La revisión sistemática de la literatura documenta tasas de falsos positivos superiores al 60 % en escritura académica producida por hablantes no nativos de inglés, incluidos estudiantes ecuatorianos y latinoamericanos. Concluimos que, cuando una institución cuenta con una reglamentación que limita el uso de IA en documentos académicos y aplica un detector como herramienta de verificación, la maestría cognitiva demostrada por el estudiante sobre dicho documento debe operar como criterio primario para resolver la ambigüedad algorítmica, dado que es la única evidencia que captura el componente de indeterminación I que el escalar del detector suprime.

Palabras clave: conjuntos neutrosóficos de valor único; detectores de IA; falsos positivos; integridad académica; Principio de Primacía Cognitiva; sesgo algorítmico; autoría académica; educación superior.

Abstract: Artificial intelligence (AI) content detectors have entered university academic assessment as instruments for verifying authorship, often without formal regulation or consideration of their documented error rates. This article proposes a single-valued neutrosophic framework (SVN) to formalize the epistemological indeterminacy that characterizes the verdicts of these systems. We argue that every AI detector collapses an implicit neutrosophic triplet $\langle T, I, F \rangle$ into a one-dimensional scalar, eliminating the indeterminacy component I and producing what we call the percentage fallacy. To resolve this indeterminacy, we propose a neutrosophic scoring function $S = T + \alpha \cdot D - I - F$ that incorporates demonstrated cognitive mastery D through oral defense or direct assessment, and we formalize the Principle of Cognitive Primacy (PCP) as the prevailing criterion for academic authorship. A systematic literature review documents false positive rates exceeding 60% in academic writing produced by non-native English speakers, including Ecuadorian and Latin American students. We conclude that when an institution has regulations limiting the use of AI in academic papers and employs a

detector as a verification tool, the student's demonstrated cognitive mastery of the document should serve as the primary criterion for resolving algorithmic ambiguity, as it is the only evidence that captures the indeterminacy component I that the detector's scalar suppresses.

Keywords: single-valued neutrosophic sets; AI detectors; false positives; academic integrity; Principle of Cognitive Primacy; algorithmic bias; academic authorship; higher education.

1. Introducción

La emergencia de los modelos de lenguaje de gran escala (LLM) en el ecosistema universitario ha desestabilizado los supuestos establecidos sobre la producción intelectual académica. Las instituciones de educación superior han respondido con velocidad desigual, desplegando detectores comerciales de contenido de IA como instrumentos de control sin marcos regulatorios que establezcan su validez como evidencia de infracción en la mayoría de los casos [4, 7]. Herramientas como GPTZero, Turnitin AI, Copyleaks y el Clasificador de OpenAI han sido incorporadas a los flujos de trabajo docentes bajo una premisa implícita que este artículo somete a escrutinio formal: que un porcentaje numérico de probabilidad de autoría artificial constituye evidencia suficiente para determinar si un estudiante utilizó IA en la elaboración de un trabajo académico. Demostramos que esta premisa es matemáticamente incorrecta.

El problema central que abordamos no es si los estudiantes usan o no la IA, sino si un escalar producido por un algoritmo constituye evidencia epistemológicamente válida para determinar la autoría. Argumentamos que la respuesta es negativa, y que este error puede formalizarse con precisión a través de la teoría de los conjuntos neutrosóficos de valor único (SVN) [8, 9], dado que la autoría académica bajo incertidumbre algorítmica es irreduciblemente un problema de indeterminación. La teoría neutrosófica es el único marco matemático que trata la indeterminación como componente independiente, no derivado de los componentes de verdad y falsedad [9, 10].

En el contexto ecuatoriano, la Ley Orgánica de Educación Superior (LOES) [19] y el Reglamento de Régimen Académico del Consejo de Educación Superior (CES) [22] no contienen ninguna disposición específica sobre el uso de detectores de IA como instrumentos de verificación de autoría. Esta ausencia no habilita ninguna práctica evaluativa: cuando las instituciones establecen normas que limitan el uso de IA en la elaboración de documentos académicos, el proceso de verificación activado debe garantizar que la evidencia empleada sea epistemológicamente válida. El principio del debido proceso exige que cualquier acción disciplinaria se sustente en evidencia suficiente, no en probabilidades estadísticas no validadas para el contexto específico de aplicación.

Este artículo realiza cinco contribuciones. Primera: reformulamos el resultado de un detector de IA como una tripleta neutrosófica de valor único $\langle T, I, F \rangle$. Segunda: definimos una función de puntuación neutrosófica $S(\text{trabajo}) = T + \alpha \cdot D - I - F$ que incorpora la maestría cognitiva demostrada D . Tercera: formulamos el Principio de Primacía Cognitiva (PPC). Cuarta: documentamos la paradoja de los falsos positivos históricos. Quinta: analizamos las implicaciones del PPC para contextos institucionales con y sin regulación específica sobre el uso de IA. Adicionalmente, proporcionamos tres tablas ilustrativas que operacionalizan el marco: un análisis de frontera de decisión (Tabla 1), un caso de inconsistencia entre detectores (Tabla 2) y una matriz de estimación de I por perfil de escritor (Tabla 3).

2. Preliminares Neutrosóficos

Presentamos las definiciones formales necesarias para el desarrollo del marco propuesto. La teoría neutrosófica fue introducida por Smarandache [9] como extensión de la lógica difusa de Zadeh [12] y de los conjuntos intuicionistas de Atanassov [11], con la diferencia fundamental de que los tres componentes de un elemento neutrosófico son matemáticamente independientes entre sí.

Definición 1 (Conjunto neutrosófico [9, 10]). Sea X un universo de discurso. Un conjunto neutrosófico A sobre X se define como:

$$A = \langle x, T_A(x), I_A(x), F_A(x) \rangle : x \in X \quad (1)$$

donde $T_A(x), I_A(x), F_A(x) : X \rightarrow]-0, 1+[$ son las funciones de membresía de verdad, indeterminación y falsedad, respectivamente, con:



$$0 \leq T_A(x) + I_A(x) + F_A(x) \leq 3 \quad (2)$$

La independencia entre T, I y F distingue esta estructura de la lógica difusa clásica, donde la condición $T + F \leq 1$ es obligatoria, y de los conjuntos intuicionistas [11], donde $T + F + H = 1$. En la teoría neutrosófica ninguna de estas restricciones aplica, lo que permite representar estados de conocimiento genuinamente indeterminados sin forzar una distribución entre verdad y falsedad.

Definición 2 (Conjunto neutrosófico de valor único — SVN [8]). Un SVN sobre X es una tripleta $A = \langle T_A, I_A, F_A \rangle$ con $T_A, I_A, F_A \in [0,1]$ y $0 \leq T_A + I_A + F_A \leq 3$ (1)

Los SVN son la restricción práctica de los conjuntos neutrosóficos a valores únicos en $[0, 1]$, lo que los hace directamente operacionales en problemas de toma de decisiones con incertidumbre real [15, 17, 21]. Cada componente se cuantifica de forma independiente, sin que el valor de uno determine o restrinja el de los demás.

Definición 3 (Función de puntuación neutrosófica [15, 16]). Sea $A = \langle T, I, F \rangle$ un SVN. La función de puntuación de A se define como:

$$s(A) = (T - I - F)/3, s(A) \in [-1,1] \quad (3)$$

Esta función penaliza la indeterminación y la falsedad, y permite ordenar alternativas neutrosóficas. Valores de s cercanos a 1 indican alta verdad y baja incertidumbre; valores cercanos a -1 indican alta falsedad o alta indeterminación.

Definición 4 (Puntuación neutrosófica extendida con maestría cognitiva). Sea W un trabajo académico evaluado por un detector de IA, con veredicto representado como $\langle T, I, F \rangle$. Sea $D \in [0, 1]$ la maestría cognitiva demostrada por el estudiante sobre W, y $\alpha \in [0, 1]$ un parámetro de ponderación institucional. La puntuación neutrosófica extendida del trabajo se define como:

$$S(\text{trabajo}) = T + \alpha \cdot D - I - F \quad (4)$$

Cuando $\alpha = 1$, la maestría cognitiva recibe el mismo peso que el componente de verdad. Cuando $\alpha = 0$, la maestría cognitiva se descarta, lo que equivale a aceptar el veredicto algorítmico sin corrección humana. Argumentamos que $\alpha = 0$ es epistemológicamente inadmisibles en cualquier sistema de evaluación académica que reconozca la existencia de falsos positivos.

Observación 1 (Justificación de la forma aditiva de S). La estructura aditiva $S = T + \alpha \cdot D - I - F$ refleja la lógica epistemológica de la verificación de autoría: D no es un término adicional al resultado del detector, sino una fuente independiente de evidencia del proceso intelectual humano, con peso institucional regulado por α . Una formulación equivalente que hace esto explícito es $T' = \lambda T + (1-\lambda)D$, $S = T' - \beta I - \gamma F$, donde la maestría cognitiva actualiza el componente de verdad en lugar de sumarse a él; ambas formas son analíticamente equivalentes bajo $\alpha = (1-\lambda)/\lambda$ y $\beta = \gamma = 1$. Conservamos la primera formulación por simplicidad analítica. La decisión institucional crítica es $\alpha > 0$, que garantiza que ningún veredicto algorítmico pueda aceptarse sin un correctivo humano basado en el proceso.

Observación 2 (Operacionalización de I). El componente de indeterminación I no es directamente observable, pero puede estimarse mediante tres métodos complementarios que las instituciones pueden combinar. (1) I_1 = tasa de falsos positivos documentada para una población comparable a los estudiantes evaluados en lengua, registro académico y género textual; esta estimación se apoya en estudios como Liang et al. [1] y Weber-Wulff et al. [2]. (2) $I_2 = 1 - r$, donde r es la confiabilidad inter-calificador validada del detector sobre una muestra de calibración local; cuando r no está disponible, I_2 debe tomar un límite superior conservador (por ejemplo, $I_2 = 0,20$) que refleje la ausencia de validación institucional. (3) $I_3 = f(\text{idoma, género, extensión, consistencia de estilo, acuerdo entre detectores})$, un factor compuesto que aumenta cuando el texto está escrito en un idioma subrepresentado en el corpus de entrenamiento del detector, cuando el género es formalmente estandarizado (jurídico, técnico, científico) o cuando dos o más detectores producen resultados contradictorios para el mismo texto. En la práctica, una institución puede establecer $I = \max(I_1, I_2)$ como límite inferior conservador y ajustar al alza en función de los factores I_3 . En el caso ilustrativo de la Sección 6, $I = 0,10$ es una estimación conservadora anclada en las tasas de falsos positivos documentadas para prosa académica en español [1, 2].

3. Estado del Arte sobre Falsos Positivos en Detectores de IA



La literatura científica publicada entre 2022 y 2025 documenta consistentemente la baja confiabilidad de los detectores de contenido de IA, con particular concentración en el fenómeno de los falsos positivos y su distribución asimétrica entre poblaciones de escritores.

Weber-Wulff et al. [2] evaluaron catorce herramientas de detección de texto de IA y concluyeron que ninguna era suficientemente precisa o confiable para su uso en contextos académicos con consecuencias disciplinarias. Su análisis reveló sesgos sistemáticos hacia la clasificación de texto humano como generado por máquina, especialmente cuando el texto presentaba estructuras sintácticas regulares o vocabulario técnico estandarizado. Elkhatat, Elsaid y Almeer [3] evaluaron cinco detectores comerciales con textos producidos por GPT-3.5 y GPT-4, y encontraron inconsistencias significativas en la clasificación de los controles humanos: los instrumentos produjeron falsos positivos y clasificaciones inciertas incluso cuando el texto fue escrito por los propios investigadores.

El hallazgo más relevante para el contexto latinoamericano proviene de Liang et al. [1], quienes demostraron mediante experimento controlado que los detectores de IA clasifican como generados por máquina más del 60 % de los ensayos TOEFL escritos por hablantes no nativos de inglés, mientras que los textos de estudiantes angloparlantes nativos son correctamente identificados como humanos en casi el 100 % de los casos. Esta brecha no refleja el uso real de IA, sino la regularidad estadística en la estructura léxica y sintáctica característica de la escritura académica formal en una segunda lengua. Los detectores confunden la competencia formal con artificialidad algorítmica.

Sadasivan et al. [6] demostraron formalmente que la detección perfecta de texto generado por LLM es, en condiciones generales, matemáticamente imposible. Su argumento establece que cualquier detector puede ser superado mediante paráfrasis simple, y que la tasa de error crece inevitablemente cuando el texto analizado proviene de escritores con estilos formales consolidados. Mitchell et al. [14], con su sistema DetectGPT, obtuvieron resultados aceptables solo en condiciones de laboratorio que no se reproducen en la práctica académica ordinaria.

La paradoja histórica representa el argumento más contundente contra la validez epistemológica de los detectores. Khalil y Er [13] documentaron que herramientas como GPTZero asignan probabilidades de origen artificial mayores que cero a textos escritos antes de la existencia de los LLM modernos. Este resultado demuestra que los detectores no identifican autoría artificial: identifican regularidad estadística lingüística. Un instrumento que asigna probabilidad de origen artificial a la prosa académica del siglo XIX no puede operar como evidencia válida de autoría en el siglo XXI.

Giray [18] documenta que las falsas acusaciones derivadas de falsos positivos han causado daños verificables a la trayectoria académica de estudiantes e investigadores, constituyendo una forma de discriminación algorítmica cuyas consecuencias son asimétricas según el idioma de escritura y el nivel de formalidad estilística. Foltýnek, Meuschke y Gipp [25] señalan que los sistemas de detección de producción no original requieren, como condición mínima de validez, tasas de error suficientemente bajas para el contexto específico de aplicación, condición que los detectores de IA actuales no satisfacen. Kasneci et al. [20] argumentan que la integración de la IA en la educación requiere marcos de evaluación que reconozcan la complejidad del aprendizaje mediado tecnológicamente, en lugar de instrumentos de control de validez no demostrada.

Cotton, Cotton y Shipway [7] argumentan que la integridad académica en la era de los LLM requiere una reconceptualización de los mecanismos de verificación, con énfasis en la evaluación del proceso sobre el producto textual. Bozkurt [5] señala que la co-autoría humano-IA desafía los marcos convencionales de autoría académica y exige criterios que vayan más allá de la inspección del artefacto final. Perkins [4] documentó que el profesorado percibe los detectores como instrumentos de control convenientes pero epistemológicamente frágiles, y señala la necesidad de mecanismos alternativos de verificación basados en la demostración directa de conocimiento.

4. La Falacia del Porcentaje: Formalización Neutrosófica

Denominamos falacia del porcentaje a la práctica de tomar el resultado escalar de un detector de IA como evidencia suficiente para determinar la autoría académica artificial. El argumento formal que desarrollamos muestra que esta práctica implica una simplificación matemática que elimina precisamente la información más relevante para una decisión justa.

Sea W un trabajo académico. Un detector de IA aplica una función $f: W \rightarrow [0, 1]$ sobre W que produce un escalar p , interpretado como la probabilidad de que W sea de origen artificial. Desde la perspectiva

neutrosófica, la autoría de W bajo incertidumbre algorítmica debe representarse como una tripleta $\langle T, I, F \rangle$ donde T es el grado de evidencia de autoría humana verificable, I es el grado de indeterminación inherente al veredicto algorítmico, y F es el grado de evidencia de autoría artificial.

Proposición 1 (Insuficiencia epistemológica del escalar de detección). El escalar de salida p de un detector de IA implica la suposición $T = 1 - F$, que corresponde a la lógica difusa clásica y no a la lógica neutrosófica. Esta suposición elimina por construcción la indeterminación I y produce una representación epistemológicamente incompleta de la autoría.

Demostración. Si el detector produce $p = F$ como probabilidad de origen artificial, y presenta $(1 - p)$ como probabilidad de origen humano, entonces $T := 1 - F$. Esta identidad colapsa la independencia entre T y F e implícitamente establece $I = 0$. Pero $I = 0$ significaría que no existe ningún grado de indeterminación en el veredicto, lo que contradice la evidencia empírica documentada [1, 2, 3, 6], que muestra clasificaciones incorrectas, contradictorias y sistemáticamente sesgadas. Por tanto, la representación escalar del detector no captura la estructura epistémica real del problema. \square

La consecuencia directa de la Proposición 1 es que ningún detector de IA, por su diseño estadístico fundamental, puede producir evidencia suficiente de autoría artificial. Puede producir indicadores de riesgo que justifiquen verificación adicional, pero no constituye prueba.

Observación 3 (Precisión técnica sobre el escalar de salida p). Representamos el resultado del detector como $p = F$ en la tripleta neutrosófica. Debe señalarse que los fabricantes comerciales no presentan este escalar como una probabilidad posterior calibrada de autoría por IA: Turnitin describe su indicador como el porcentaje de texto que su modelo determina que "podría" haber sido generado o modificado por herramientas de IA, y señala explícitamente que el indicador no debe servir como única base para decisiones disciplinarias [Turnitin AI Writing Detection, 2024]. OpenAI retiró su propio clasificador precisamente por precisión insuficiente [23]. A pesar de estas interpretaciones variables, cualquier salida escalar que carezca de un componente explícito de indeterminación suprime I por construcción, que es la afirmación epistemológica de la Proposición 1. La formalización como $p = F$ es, por tanto, una abstracción fundamentada que hace explícita la pérdida de información, con independencia de la calibración específica del instrumento de cada fabricante.

Proposición 2 (Expresividad diferencial de la puntuación extendida). Sea $\langle T, I, F \rangle$ la representación neutrosófica del veredicto de un detector de IA con $I > 0$, y sean $D \in (0, 1]$ y $\alpha \in (0, 1]$. La puntuación extendida $S(\text{trabajo}) = T + \alpha \cdot D - I - F$ distingue formalmente dos situaciones evaluativas que el criterio escalar $p = F$ trata como idénticas: (a) cuando $T + \alpha \cdot D > I$, la puntuación $S > 0$ mientras que el criterio escalar $-F \leq 0$, apoyando un hallazgo de autoría verificada; (b) cuando $T + \alpha \cdot D \leq I$, la puntuación $S \leq 0$, indicando que la indeterminación y la evidencia de origen artificial predominan a pesar de la demostración parcial de maestría cognitiva.

Demostración. El criterio escalar correspondiente exclusivamente al veredicto del detector es $-F \in [-1, 0]$. La diferencia entre la puntuación extendida y este criterio es:

$$\Delta = S - (-F) = T + \alpha \cdot D - I. \quad (5)$$

Caso (a): Si $T + \alpha \cdot D > I$, entonces $\Delta > 0$, de modo que $S > -F$. Dado que $F \in [0, 1]$ implica $-F \leq 0$ mientras que $S > 0$, la puntuación extendida apoya un hallazgo de autoría verificada donde el criterio escalar no lo hace.

Caso (b): Si $T + \alpha \cdot D \leq I$, entonces $\Delta \leq 0$, de modo que $S \leq 0$, confirmando que la indeterminación y la evidencia de autoría artificial prevalecen con independencia del resultado específico F del detector.

El Caso (a) es el escenario esperado para estudiantes que genuinamente elaboraron su trabajo: la defensa oral directa permite que D se aproxime a 1 (véase Sección 6), mientras que I , acotada por las tasas de falsos positivos documentadas [1, 2, 3], típicamente no supera 0,15–0,20 para textos individuales, de modo que $T + \alpha \cdot D > I$ se cumple para cualquier $\alpha > 0$. El Caso (b) aplica cuando la maestría cognitiva demostrada es insuficiente para compensar el peso combinado de la indeterminación y la evidencia de falsedad, lo que es consistente con el escenario de un estudiante que no puede defender el contenido de su propio trabajo. \square

5. El Principio de Primacía Cognitiva (PPC): Marco Formal

Definición 5 (Maestría cognitiva). Sea W un trabajo académico y s un estudiante. La maestría cognitiva $D(s, W) \in [0, 1]$ de s sobre W es una medida que cuantifica la capacidad de s para responder, explicar, ampliar y relacionar el contenido de W bajo condiciones de evaluación directa, mediante un protocolo de defensa oral estructurado y contradictorio conducido por un evaluador humano.

La medida D incorpora al menos tres dimensiones verificables: la capacidad de reformular las ideas centrales del trabajo sin consultar el texto; la capacidad de responder preguntas en profundidad sobre el tema



abordado; y la capacidad de identificar las fuentes primarias citadas y explicar su relevancia. Un evaluador capacitado puede asignar D mediante una rúbrica explícita, lo que hace al PPC operacionalizable en contextos institucionales concretos.

Principio de Primacía Cognitiva (PPC). Sea W un trabajo académico sobre el cual un detector de IA ha producido un veredicto $\langle T_d, I_d, F_d \rangle$ con $I_d > 0$. Sea $D(s, W)$ la maestría cognitiva demostrada por el estudiante s en la defensa. El PPC establece:

$$SiS(trabajo) = T_d + \alpha \cdot D(s, W) - I_d - F_d > 0, \quad (6)$$

entonces W debe considerarse un trabajo de autoría humana verificada, con independencia del valor de F_d . En consecuencia, ningún proceso de verificación sobre el uso de IA puede resolverse exclusivamente con el escalar $p = F_d$.

El fundamento epistemológico del PPC descansa en la distinción entre evidencia de proceso y evidencia de producto. Un detector de IA examina únicamente el artefacto textual final. Pero la autoría académica es un atributo del proceso intelectual que generó ese artefacto, no del artefacto en sí. Un texto puede presentar características estadísticas similares a las del texto generado por IA porque el estudiante escribe en una lengua que no es su lengua materna [1], porque ha interiorizado profundamente los patrones del discurso académico formal [18], o porque su estilo es técnicamente riguroso [2]. Ninguna de estas causas implica uso de IA. Solo la evaluación directa del proceso cognitivo resuelve esta indeterminación.

El PPC es también consistente con el principio del debido proceso: la carga de la prueba recae sobre quien acusa, y el acusado tiene derecho a presentar evidencia exculpatoria. La demostración positiva de maestría cognitiva es precisamente esa evidencia, y su peso epistemológico es superior al de un porcentaje estadístico cuyos errores de Tipo I superan el 60 % en poblaciones comparables a la ecuatoriana [1].

6. Caso Numérico Ilustrativo

Presentamos un caso hipotético que ilustra la aplicación del marco neutrosófico con valores numéricos concretos e internamente verificables.

Sean s_1 y s_2 dos estudiantes que presentan trabajos académicos W_1 y W_2 respectivamente. Ambos obtienen el mismo porcentaje del detector de IA: $p = 0,65$ (65 % de probabilidad de origen artificial). Bajo la evaluación escalar convencional, ambos estudiantes recibirían una respuesta institucional idéntica.

Paso 1: Representación neutrosófica del veredicto del detector.

Dado que el detector solo produce el escalar $p = F$, e implícitamente asume $T = 1 - F$, las tripletas iniciales son idénticas. Corrigiendo con $I > 0$ según las tasas de error documentadas [2, 3]:

$$W_1: \langle T_1, I_1, F_1 \rangle = \langle 0.25, 0.10, 0.65 \rangle$$

$$W_2: \langle T_2, I_2, F_2 \rangle = \langle 0.25, 0.10, 0.65 \rangle$$

Paso 2: Defensa oral y cuantificación de la maestría cognitiva.

$$D(s_1, W_1) = 0,90 \quad (\text{maestría alta, respuestas coherentes y autónomas})$$

$$D(s_2, W_2) = 0,15 \quad (\text{maestría baja, respuestas superficiales e inconsistentes})$$

Paso 3: Cálculo de la puntuación neutrosófica extendida con $\alpha = 1$.

$$S(W_1) = 0,25 + 1 \times 0,90 - 0,10 - 0,65 = 0,40 > 0 \rightarrow \text{autoría verificada}$$

$$S(W_2) = 0,25 + 1 \times 0,15 - 0,10 - 0,65 = -0,35 < 0 \rightarrow \text{requiere investigación adicional}$$

s_1 obtiene una puntuación positiva que, por el PPC, indica autoría humana verificada a pesar del porcentaje del detector. s_2 obtiene una puntuación negativa que genera un caso que requiere investigación adicional. Este ejemplo numérico demuestra que dos estudiantes con el mismo porcentaje de detección de IA pueden tener situaciones radicalmente diferentes respecto a su autoría, y que solo el marco SVN con la variable D puede capturar formalmente esta diferencia.

Para ilustrar la paradoja histórica, considérese un texto académico H escrito en 1994. Si se aplica un detector contemporáneo a H , el resultado sería $F_H > 0$ a pesar de que H fue producido antes de la existencia de cualquier LLM. La representación neutrosófica requiere $I_H > 0$, lo que activa inmediatamente el PPC: el veredicto es indeterminado y no puede usarse como evidencia de infracción.



La Tabla 1 generaliza el caso numérico calculando, para cada nivel de resultado del detector F, el umbral crítico de maestría cognitiva D^* por debajo del cual el PPC no puede certificar la autoría humana. D^* se deriva analíticamente de $S(\text{trabajo}) = 0$, con $T = 1 - F$ (la suposición escalar), $I = 0,10$ (estimación conservadora de falso positivo [1, 2]) y $\alpha = 1$: $D^* = \max(0, F + I - T) = \max(0, 2F - 0,90)$. La tabla muestra que solo para $F \geq 0,55$ el marco requiere maestría cognitiva no trivial, y solo para $F \geq 0,85$ el D^* requerido supera 0,80, condiciones que un autor genuino debería satisfacer sin dificultad.

Tabla 1. Umbral crítico de maestría cognitiva D^* para autoría verificada en función del resultado del detector de IA F ($T = 1 - F$; $I = 0,10$; $\alpha = 1$).

F (resultado del detector)	T (implícito por el escalar)	I (estimación conservadora)	Umbral D^*	Resultado institucional
0.40	0.60	0.10	0.00 (cualquier D es suficiente)	Autoría verificada sin defensa
0.55	0.45	0.10	0.20	Defensa mínima requerida
0.65	0.35	0.10	0.40	Defensa moderada requerida
0.75	0.25	0.10	0.60	Defensa sólida requerida
0.85	0.15	0.10	0.80	Defensa extensa requerida
0.95	0.05	0.10	1.00	Defensa exhaustiva requerida

Nota. $D^* = \max(0, F + I - T)$. Los valores asumen la estimación conservadora de I derivada de Liang et al. [1] y Weber-Wulff et al. [2]. Valores más altos de I (véase Observación 2) desplazan D^* hacia abajo, facilitando la certificación; $\alpha < 1$ tiene el mismo efecto.

Un segundo caso ilustrativo proporciona sustento empírico para $I > 0$. Considérese un ensayo académico en español W_3 producido por un hablante no nativo de inglés y evaluado por tres detectores comerciales. La Tabla 2 reporta las puntuaciones de IA obtenidas para el mismo texto por GPTZero, Turnitin AI y Copyleaks, consistentes con los rangos de rendimiento documentados por Weber-Wulff et al. [2] y Elkhatat et al. [3] para corpora académicos. La desviación típica de 0,17 entre detectores es en sí misma evidencia directa de que $I > 0$: si la indeterminación fuera cero, todos los instrumentos que miden la misma propiedad convergerían. La divergencia hace que $I_{\min} \geq \sigma = 0,17$, validando el $I = 0,10$ conservador utilizado en la Tabla 1 como límite inferior y justificando $I_2 = 1 - r$ en la Observación 2.

Tabla 2. Evaluación ilustrativa entre detectores del ensayo académico en español W_3 . Las puntuaciones de IA siguen los rangos documentados en la literatura [2, 3] para escritura académica de hablantes no nativos.

Detector	Puntuación IA (F)	Puntuación humana (1-F)	Veredicto	Fuente
GPTZero	0.72	0.28	Probablemente generado por IA	Weber-Wulff et al. [2]
Turnitin AI	0.38	0.62	Incierto	Weber-Wulff et al. [2]
Copyleaks	0.61	0.39	Probablemente generado por IA	Elkhatat et al. [3]
Media	0.57	0.43	—	—
Desv. típ.	0.17	0.17	—	—

Nota. Las puntuaciones son valores ilustrativos consistentes con los rangos de rendimiento reportados. La desviación típica inter-detector $\sigma = 0,17$ proporciona un límite inferior empírico para I en la representación neutrosófica de W_3 . Bajo el criterio escalar, los tres veredictos desencadenarían una respuesta institucional idéntica a pesar de su desacuerdo sustantivo.

7. Discusión



7.1. Confiabilidad de los Detectores de IA y la Falacia del Porcentaje

El análisis formalizado en las Proposiciones 1 y 2 muestra que todo detector de IA que produce un escalar como salida elimina estructuralmente la indeterminación neutrosófica, lo que invalida su uso como evidencia suficiente de autoría artificial. Esto es consistente con lo señalado por Weber-Wulff et al. [2], quienes documentan que ninguna de las catorce herramientas de detección evaluadas en su estudio alcanzó niveles suficientes de precisión o confiabilidad para contextos académicos con consecuencias disciplinarias. Según estos autores, los detectores presentan sesgos sistemáticos que los hacen inadecuados como instrumentos de verificación de autoría, lo que confirma que el escalar de salida de estos sistemas no puede operar como evidencia concluyente por sí solo.

En la misma línea, Elkhatat, Elsaid y Almeer [3] encontraron que cinco detectores comerciales producen clasificaciones inconsistentes incluso cuando el texto fue escrito íntegramente por los propios investigadores, generando falsos positivos en los controles humanos. Esto coincide con la Proposición 1: si los detectores clasifican como artificial un texto producido por evaluadores académicos, el componente de indeterminación I es estructuralmente mayor que cero, lo que colapsa cualquier interpretación escalar del resultado. A diferencia de la práctica implícita en algunas directrices institucionales que citan el porcentaje de Turnitin o GPTZero como criterio de verificación directo, el presente análisis muestra que esta práctica elimina, por construcción matemática, la información más relevante para una decisión justa.

Sadasivan et al. [6] establecen, mediante un argumento formal, que la detección perfecta de texto generado por LLM es matemáticamente imposible en condiciones generales, dado que cualquier detector puede ser superado mediante paráfrasis simple y la tasa de error crece estructuralmente cuando el texto proviene de escritores con estilos consolidados. Esto es consistente con la Proposición 2: la puntuación extendida S supera al criterio escalar precisamente en los casos donde el estilo formal del estudiante eleva artificialmente el valor de F . Por otro lado, Mitchell et al. [14] argumentan que su sistema DetectGPT alcanza un rendimiento aceptable en condiciones controladas de laboratorio; sin embargo, los propios autores reconocen que estas condiciones no se reproducen en la práctica académica ordinaria, por lo que la capacidad discriminadora real del detector sigue siendo insuficiente para fundamentar decisiones con consecuencias en la trayectoria académica del estudiante.

La paradoja histórica documentada por Khalil y Er [13], quienes observaron que detectores como GPTZero asignan probabilidades de origen artificial mayores que cero a textos producidos antes de la existencia de los LLM modernos, refuerza este argumento desde un ángulo empírico. Según estos autores, el fenómeno evidencia que los detectores no identifican autoría artificial sino regularidad estadística lingüística, lo que converge con la tesis central de este artículo: el instrumento detecta patrones de frecuencia, no procesos cognitivos, y esta limitación es irreducible en el diseño actual de estos sistemas.

7.2. Sesgo Lingüístico y Discriminación Algorítmica en el Contexto Latinoamericano

El análisis revela que los detectores de IA penalizan estructuralmente a quienes escriben en idiomas distintos al inglés o en estilos académicos formales. Esto concuerda directamente con los hallazgos de Liang et al. [1], quienes documentaron que más del 60 % de los ensayos TOEFL escritos por hablantes no nativos de inglés fueron clasificados como generados por IA, frente a tasas de identificación humana correcta cercanas al 100 % para textos en inglés nativo. Según estos autores, la brecha no refleja el uso real de IA sino la regularidad estadística léxico-sintáctica propia del registro académico formal en una segunda lengua, lo que convierte al detector en un instrumento que discrimina según la competencia lingüística, no según la autoría.

En el contexto ecuatoriano, donde los trabajos académicos se producen en español y los estudiantes construyen su escritura dentro de las convenciones del discurso académico latinoamericano, el fenómeno descrito por Liang et al. [1] se traduce en tasas de falsos positivos potencialmente superiores a las documentadas para el inglés, dado que los detectores comerciales se entrenan predominantemente con corpora en ese idioma. Esto es consistente con lo que señala Giray [18], quien documenta que los investigadores con estilos de escritura formalmente rigurosos obtienen recurrentemente tasas de detección falsa más altas. Según Giray [18], las falsas acusaciones derivadas de estos errores han generado daños verificables en la trayectoria de estudiantes y académicos, lo que refuerza la necesidad del PPC como mecanismo que permite corregir ese sesgo mediante la medición directa de la maestría cognitiva.

La Tabla 3 operacionaliza la Observación 2 estimando I para cuatro perfiles de escritor representativos del contexto universitario latinoamericano. I_1 está anclada a las tasas de falsos positivos documentadas [1, 2, 3]; $I_2 = 1 - r$ utiliza $r = 0,85$ como estimación conservadora de confiabilidad inter-calificador para detectores



comerciales [2]; I_3 refleja los factores lingüísticos y genéricos compuestos. El I adoptado = $\max(I_1, I_2)$ establece la indeterminación mínima que cualquier proceso de evaluación institucional debe reconocer antes de aplicar el PPC.

Tabla 3. Indeterminación I estimada por perfil de escritor aplicando la operacionalización de tres métodos de la Observación 2 ($r = 0,85$ estimación conservadora de confiabilidad [2]).

Perfil del escritor	I_1 (tasa FP, lit.)	$I_2 = 1 - r$	I_3 (compuesto)	I adoptado	Fuente
Inglés nativo, texto informal	0.02	0.15	Bajo	0.15	Liang et al. [1]
Inglés no nativo, ensayo académico	0.61	0.15	Alto	0.61	Liang et al. [1]
Español (Ecuador), prosa técnica	≥ 0.30 est.	0.15	Alto	0.30	[1, 2, 3]
Texto histórico (anterior a 2020)	N/A	1.00	Máximo	0.60	Khalil & Er [13]

Nota. I_1 para prosa española/ecuatoriana se estima como límite inferior: ningún estudio hasta la fecha ha medido tasas de falsos positivos específicamente para escritura académica en español L1. El valor $\geq 0,30$ se deriva por extrapolación descendente de Liang et al. [1], dado que los detectores comerciales se entrenan predominantemente con corpora en inglés. I adoptado = $\max(I_1, I_2)$.

A diferencia de la lectura que considera a los detectores de IA como herramientas de apoyo docente neutrales, lectura subyacente en algunos documentos de política educativa institucional, el presente análisis muestra que la neutralidad algorítmica es una ilusión cuando el instrumento fue diseñado y validado sobre un corpus lingüístico que no representa a la población que lo recibe. Cotton, Cotton y Shipway [7] reconocen que los detectores pueden servir como señal de alerta inicial, pero argumentan que la integridad académica en la era de los LLM requiere privilegiar la evaluación del proceso intelectual sobre la inspección del artefacto textual. Esto concuerda con el PPC: la maestría cognitiva demostrada D es precisamente esa evidencia de proceso que los autores señalan como criterio superior, y su incorporación formal en la función S le otorga el peso operativo que la evaluación equitativa requiere.

7.3. La Maestría Cognitiva como Criterio de Verificación bajo Regulación Institucional del Uso de IA

El argumento central de este artículo no busca crear mecanismos sancionatorios para quienes usan IA, sino establecer qué debe verificarse cuando una institución ya cuenta con reglamentación que limita o regula ese uso y aplica un detector como herramienta de verificación. En ese escenario, el problema es preciso: si el detector emite un veredicto ambiguo sobre un documento presuntamente generado con IA, ¿qué evidencia puede resolver esa ambigüedad? Esto concuerda con lo que plantea Perkins [4], quien señala que la ausencia de criterios claros sobre qué constituye evidencia válida de infracción transfiere a la discrecionalidad docente una decisión de alto impacto sobre la trayectoria del estudiante. Según este autor, esa discrecionalidad genera inequidades sistemáticas que afectan desproporcionalmente a los estudiantes con menor capacidad institucional para impugnar.

Bozkurt [5] argumenta que la irrupción de la IA generativa obliga a reconsiderar conceptos como la autoría y la responsabilidad académica, y que las instituciones deben adoptar marcos que reconozcan la naturaleza colaborativa de la producción contemporánea. Si bien el presente análisis coincide con Bozkurt [5] en que los marcos convencionales de autoría son insuficientes para el contexto actual, difiere en el enfoque: mientras ese autor orienta la discusión hacia la redefinición de las normas de autoría compartida, el PPC propuesto no debate si el uso de IA es legítimo o no, sino que establece el criterio operativo para determinar cuándo el estudiante demuestra que el conocimiento contenido en el documento le pertenece, con independencia de las herramientas utilizadas durante la elaboración. La demostración positiva de un $D(s,W)$ alto constituye la evidencia de autoría intelectual que ningún porcentaje algorítmico puede proporcionar.

Kasneci et al. [20] argumentan que la integración de la IA en la educación requiere marcos que equilibren oportunidades pedagógicas con salvaguardas éticas, y advierten que las respuestas institucionales puramente restrictivas pueden obstaculizar el desarrollo de competencias que los estudiantes necesitarán en su vida profesional. Esto concuerda con el argumento del PPC: verificar la maestría cognitiva no desincentiva el uso reflexivo de la IA en el proceso de aprendizaje, sino que traslada la responsabilidad académica al lugar correcto. Cuando el estudiante puede sostener una defensa sólida sobre el contenido de su trabajo, con independencia de qué herramientas utilizó durante la elaboración, el sistema de evaluación cumple su propósito pedagógico de verificar la apropiación real del conocimiento.

Zawacki-Richter et al. [27] documentan que las aplicaciones de IA en educación superior han sido adoptadas fundamentalmente desde la perspectiva tecnológica, con escasa participación docente en el diseño de los criterios de validación. Esto concuerda con la crítica subyacente de este artículo: los detectores de IA ingresaron a los procesos de evaluación universitaria sin validación pedagógica específica para cada contexto de uso, reproduciendo el patrón de primacía tecnológica sobre el criterio educativo que esos autores identifican como riesgo estructural. La función $S = T + \alpha \cdot D - I - F$ invierte esta lógica al situar el parámetro α , decidido por la institución en función de sus objetivos pedagógicos, como regulador del peso relativo entre el veredicto algorítmico y la evidencia humana directa.

Foltýnek, Meuschke y Gipp [25] establecen que los sistemas de detección de producción no original requieren, como condición mínima de validez en contextos académicos con consecuencias disciplinarias, tasas de error suficientemente bajas para la población específica de uso. Esta condición es equivalente, en términos neutrosóficos, a exigir que I sea significativamente cercano a cero antes de aceptar el veredicto como evidencia. Dado que los estudios de Liang et al. [1], Weber-Wulff et al. [2] y Elkhatat et al. [3] demuestran convergentemente que esta condición no se cumple para ninguna herramienta disponible en el contexto latinoamericano, la conclusión es directa: cuando una institución cuenta con reglamentación que limita el uso de IA y decide verificar la autoría de un documento cuestionado, el PPC debe operar como criterio obligatorio de resolución, dado que convierte la indeterminación I del detector en una pregunta empíricamente responsable mediante evaluación humana directa.

8. Conclusiones

Este artículo ha demostrado mediante formalización neutrosófica que el resultado de todo detector de IA colapsa una tripleta neutrosófica implícita $\langle T, I, F \rangle$ en un escalar unidimensional, eliminando la indeterminación I por diseño. Esta eliminación es epistemológicamente incorrecta porque la evidencia empírica documenta consistentemente que $I > 0$ para cualquier texto académico evaluado con las herramientas actualmente disponibles.

El Principio de Primacía Cognitiva que proponemos resuelve formalmente esta indeterminación incorporando la maestría cognitiva demostrada D como variable decisiva central a través de la función de puntuación extendida $S = T + \alpha \cdot D - I - F$. Este principio no propone nuevas sanciones por el uso de IA: establece qué debe verificarse cuando una institución ya cuenta con reglamentación que limita ese uso y emplea un detector como herramienta de verificación. La respuesta es que la maestría cognitiva demostrada por el estudiante sobre el documento cuestionado —y no el porcentaje algorítmico— debe operar como criterio primario de resolución.

La revisión sistemática de la literatura confirma tres hechos: los detectores de IA presentan tasas de falsos positivos superiores al 60 % para la escritura de hablantes no nativos de inglés; estas herramientas asignan probabilidades de origen artificial a textos históricos producidos antes de la existencia de los LLM; y ninguna herramienta disponible alcanza niveles de precisión suficientes para su uso en procesos con consecuencias disciplinarias.

En el contexto ecuatoriano, la ausencia de reglamentación específica del CES sobre el uso de detectores de IA implica que cualquier sanción fundamentada exclusivamente en porcentajes algorítmicos carece de base normativa adecuada. El PPC ofrece el criterio que las instituciones pueden adoptar para garantizar el debido proceso: el veredicto de un detector identifica un caso que amerita verificación, pero es la defensa oral directa la que resuelve la indeterminación.

Como líneas de trabajo futuro, proponemos extender el marco a evaluaciones grupales mediante operadores de agregación SVN [15, 17], el desarrollo de protocolos de defensa oral estandarizados con métricas formales de maestría cognitiva calibradas para contextos universitarios latinoamericanos, y la validación empírica del parámetro α mediante estudios de caso en universidades ecuatorianas.



Referencias

- [1] Liang, W.; Yuksekgonul, M.; Mao, Y.; Wu, E.; Zou, J. GPT detectors are biased against non-native English writers. *Patterns* 2023, 4(7), 100779. <https://doi.org/10.1016/j.patter.2023.100779>
- [2] Weber-Wulff, D.; Anohina-Naumeca, A.; Bjelobaba, S.; Foltýnek, T.; Guerrero-Dib, J.; Pokorný, O.; Šigut, P.; Waddington, L. Testing of detection tools for AI-generated text. *International Journal for Educational Integrity* 2023, 19(1), 26. <https://doi.org/10.1007/s40979-023-00146-z>
- [3] Elkhatat, A. M.; Elsaid, K.; Almeer, S. Evaluating the efficacy of AI content detection tools in differentiating between human and AI-generated text. *International Journal for Educational Integrity* 2023, 19(1), 17. <https://doi.org/10.1007/s40979-023-00140-5>
- [4] Perkins, M. Academic Integrity considerations of AI Large Language Models in the post-pandemic era: ChatGPT and beyond. *Journal of University Teaching and Learning Practice* 2023, 20(2), 7. <https://doi.org/10.53761/1.20.02.07>
- [5] Bozkurt, A. GenAI et al.: Cocreación, Autorship, Ownership, Academic Ethics and Integrity in a Time of Generative AI. *Open Praxis* 2024, 16(1), 1–10. <https://doi.org/10.55982/openpraxis.16.1.654>
- [6] Sadasivan, V. S.; Kumar, A.; Balasubramanian, S.; Wang, W.; Feizi, S. Can AI-generated text be reliably detected? *arXiv:2303.11156* 2023. <https://doi.org/10.48550/arXiv.2303.11156>
- [7] Cotton, D. R. E.; Cotton, P. A.; Shipway, J. R. Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International* 2024, 61(2), 228–239. <https://doi.org/10.1080/14703297.2023.2190148>
- [8] Wang, H.; Smarandache, F.; Zhang, Y. Q.; Sunderraman, R. Single valued neutrosophic sets. *Multispace and Multistructure* 2010, 4, 410–413.
- [9] Smarandache, F. *Neutrosophy: Neutrosophic Probability, Set, and Logic*. American Research Press: Rehoboth, NM, USA, 1998.
- [10] Smarandache, F. *A Unifying Field in Logics: Neutrosophic Logic*. *Neutrosophy, Neutrosophic Set, Neutrosophic Probability and Statistics*, 4th ed.; American Research Press: Rehoboth, NM, USA, 2005.
- [11] Atanassov, K. T. Intuitionistic fuzzy sets. *Fuzzy Sets and Systems* 1986, 20(1), 87–96. [https://doi.org/10.1016/S0165-0114\(86\)80034-3](https://doi.org/10.1016/S0165-0114(86)80034-3)
- [12] Zadeh, L. A. Fuzzy sets. *Information and Control* 1965, 8(3), 338–353. [https://doi.org/10.1016/S0019-9958\(65\)90241-X](https://doi.org/10.1016/S0019-9958(65)90241-X)
- [13] Khalil, M.; Er, E. Will ChatGPT get you caught? Rethinking of plagiarism detection. In *Learning and Collaboration Technologies. HCT 2023. Lecture Notes in Computer Science*, vol. 14040; Zaphiris, P., Ioannou, A., Eds.; Springer: Cham, Switzerland, 2023; pp. 475–487. https://doi.org/10.1007/978-3-031-34411-4_32
- [14] Mitchell, E.; Lee, Y.; Khazatsky, A.; Manning, C. D.; Finn, C. DetectGPT: Zero-shot machine-generated text detection using probability curvature. In *Proceedings of the 40th International Conference on Machine Learning (ICML 2023)*, PMLR 202, 24950–24962, 2023.
- [15] Ye, J. A multicriteria decision-making method using aggregation operators for simplified neutrosophic sets. *Journal of Intelligent and Fuzzy Systems* 2014, 26(5), 2459–2466. <https://doi.org/10.3233/IFS-130916>
- [16] Ye, J. Single valued neutrosophic cross-entropy for multicriteria decision making problems. *Applied Mathematical Modelling* 2014, 38(3), 1170–1175. <https://doi.org/10.1016/j.apm.2013.07.020>
- [17] Biswas, P.; Pramanik, S.; Giri, B. C. TOPSIS method for multi-attribute group decision-making under single-valued neutrosophic environment. *Neural Computing and Applications* 2016, 27(3), 727–737. <https://doi.org/10.1007/s00521-015-1891-2>
- [18] Giray, L. The problem with false positives: AI detection unfairly accuses scholars of AI plagiarism. *The Serials Librarian* 2024, 85(5–6), 181–189. <https://doi.org/10.1080/0361526X.2024.2433256>
- [19] Asamblea Nacional del Ecuador. Ley Orgánica de Educación Superior (LOES). Registro Oficial Suplemento 297, August 2, 2018.



- [20] Kasneci, E.; Seßler, K.; Küchemann, S.; Bannert, M.; Dementieva, D.; Fischer, F.; Gasser, U.; Groh, G.; Günemann, S.; Hüllermeier, E.; Krusche, S.; Kutyniok, G.; Michaeli, T.; Nerdel, C.; Pfeffer, J.; Poquet, O.; Sailer, M.; Schmidt, A.; Seidel, T.; et al. ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences* 2023, 103, 102274. <https://doi.org/10.1016/j.lindif.2023.102274>
- [21] Hechavarría-Hernández, J. R.; Leyva Vázquez, M. Y.; Smarandache, F. Neutrosophic stance detection and fsQCA-based necessary condition analysis for causal hypothesis assessment in AI-enhanced learning. *Neutrosophic Computing and Machine Learning* 2025, 40.
- [22] Consejo de Educación Superior del Ecuador (CES). Reglamento de Régimen Académico. Resolución RPC-SO-08-No.111-2019. CES: Quito, Ecuador, 2019.
- [23] OpenAI. GPT-4 Technical Report. arXiv:2303.08774 2023. <https://doi.org/10.48550/arXiv.2303.08774>
- [24] Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A. N.; Kaiser, Ł.; Polosukhin, I. Attention is all you need. *Advances in Neural Information Processing Systems (NeurIPS 2017)* 2017, 30.
- [25] Foltýnek, T.; Meuschke, N.; Gipp, B. Academic plagiarism detection: A systematic literature review. *ACM Computing Surveys* 2019, 52(6), 112. <https://doi.org/10.1145/3345317>
- [26] Brown, T.; Mann, B.; Ryder, N.; Subbiah, M.; Kaplan, J. D.; Dhariwal, P.; Neelakantan, A.; Shyam, P.; Sastry, G.; Askell, A.; Agarwal, S.; Herbert-Voss, A.; Krueger, G.; Henighan, T.; Child, R.; Ramesh, A.; Ziegler, D.; Wu, J.; Winter, C.; et al. Language models are few-shot learners. *Advances in Neural Information Processing Systems (NeurIPS 2020)* 2020, 33, 1877–1901.
- [27] Zawacki-Richter, O.; Marín, V. I.; Bond, M.; Gouverneur, F. Systematic review of research on artificial intelligence applications in higher education — where are the educators? *International Journal of Educational Technology in Higher Education* 2019, 16(1), 39. <https://doi.org/10.1186/s41239-019-0171-0>
- [28] Gehrmann, S.; Strobel, H.; Rush, A. M. GLTR: Statistical detection and visualization of generated text. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics: System Demonstrations (ACL 2019)*; Association for Computational Linguistics: Florence, Italy, 2019; pp. 111–116. <https://doi.org/10.18653/v1/P19-3019>

Recibido: Mes Día, Año. Aceptado: Mes Día, Año

