

Open Minds

Internacional Journal

ISSN 2675-5157

vol. 1, n. 2, 2025

••• ARTICLE 2

Acceptance date: 02/10/2025

PRUEBA DE CONCEPTO DE GEMELO DIGITAL ÁULICO IMPLEMENTADO A PARTIR DE ANALÍTICA EN CLASES VIRTUALES SINCRÓNICAS

Marcelo Dante Caiafa

Departamento de Ingeniería e Investigaciones Tecnológicas, Universidad Nacional de La Matanza

Alejandro Bevilacqua

Departamento de Ingeniería e Investigaciones Tecnológicas, Universidad Nacional de La Matanza

Fernando Javier Dufour

Departamento de Ingeniería e Investigaciones Tecnológicas, Universidad Nacional de La Matanza

Gustavo Rosales

Departamento de Ingeniería e Investigaciones Tecnológicas, Universidad Nacional de La Matanza



Todo o conteúdo desta revista está licenciado sob a Licença Creative Commons Atribuição 4.0 Internacional (CC BY 4.0).

Resumen: La analítica de espacios virtuales de aprendizaje se ha convertido en una poderosa e interesante herramienta para comprender y optimizar la experiencia educativa en plataformas digitales. En el presente trabajo se pretende diseñar y validar una representación digital de un aula virtual, creada usando datos reales obtenidos de clases sincrónicas en línea. Esta “prueba de concepto” permite explorar un modelo digital que contribuya a la comprensión, evaluación o incluso predicción del comportamiento educativo dentro de ese entorno. El trabajo se enfoca en el análisis y la interpretación de datos generados por los estudiantes durante su interacción en ambientes virtuales de aprendizaje proporcionando información sobre su comportamiento, desempeño y necesidades. Los datos son recolectados a partir de los registros de las grabaciones de las clases virtuales dictadas en modalidad sincrónica durante clases de Física para alumnos de ingeniería. El estudio se basa en la utilización de algoritmos de procesamiento de lenguaje natural para analizar las transcripciones de los registros grabados durante las clases y la aplicación de modelos de aprendizaje automático para calcular los principales indicadores. Las interacciones en un entorno virtual sincrónico de aprendizaje se analizan mediante la identificación de variables en la dimensión conductual, emocional y cognitiva. El presente estudio está destinado a docentes, coordinadores académicos, investigadores en educación y diseñadores instruccionales interesados en optimizar la efectividad de la enseñanza virtual. Asimismo, puede resultar útil para autoridades universitarias y responsables de programas de formación que buscan implementar modelos basados en datos objetivos.

Palabras clave: Analítica de aprendizaje, Aprendizaje automático, Procesamiento

de lenguaje natural, Ambiente virtual de aprendizaje, LLM, Gemelo Digital Áulico.

Introducción

La inteligencia artificial (IA) ha avanzado significativamente desde que John McCarthy fundó el primer laboratorio en la universidad de Stanford en 1956 [1]. El aprendizaje automático, conocido por sus siglas en inglés como ML (machine learning), es un subcampo de la IA que implica el uso de algoritmos y modelos estadísticos para permitir que las máquinas aprendan de los datos, dándole capacidad para hacer predicciones, detectar patrones y tomar decisiones sin necesidad de estar programadas explícitamente.

El procesamiento de lenguaje natural, conocido por sus siglas en inglés como NLP (Natural Language Processing), es un tipo de tecnología de machine learning que brinda a las computadoras la capacidad de interpretar, manipular y comprender el lenguaje humano. Para ello, los modelos de lenguaje (Language Models, LM) han emergido como herramientas clave en NLP, permitiendo a las máquinas comprender contextos complejos y generar textos coherentes en diferentes aplicaciones. Estos modelos aprenden patrones del idioma a partir de grandes corpus de datos, facilitando tareas como traducción automática, análisis de sentimientos, y respuestas automáticas [2].

Uno de los modelos de lenguaje más influyentes en la actualidad es BERT (Bidirectional Encoder Representations from Transformers), desarrollado por Google AI. BERT se distingue por su enfoque bidireccional, que permite comprender el contexto completo de una palabra analizando simultáneamente las palabras antes y después en

la oración, a diferencia de modelos unidireccionales que solo consideran una dirección. Además, utiliza la arquitectura Transformer, basada en mecanismos de atención, para captar relaciones a largo plazo en el texto y producir representaciones contextuales robustas [3]. Estas capacidades permiten que BERT tenga un rendimiento sobresaliente en tareas de comprensión del lenguaje y han abierto nuevas posibilidades en aplicaciones de educación, análisis de textos y sistemas de interacción automatizados.

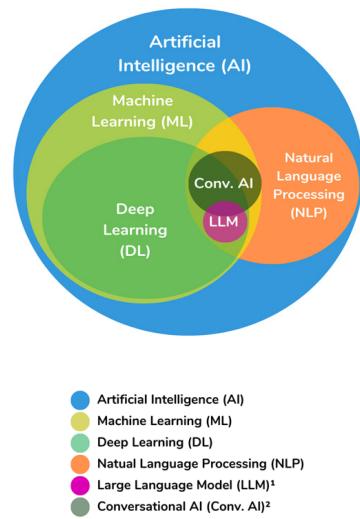
La analítica de aprendizaje y la IA han crecido junto al desarrollo de algoritmos para hacer recomendaciones, tomar decisiones y aprender en diferentes contextos. Estas diferentes técnicas emergen como herramientas para mejorar la calidad educativa de los ambientes virtuales de aprendizaje [4]. El desempeño en las clases en línea adquiere particular relevancia en los desarrollos de clases sincrónicas virtuales. La disponibilidad de datos posibilita el uso de técnicas analíticas de IA para evaluar distintos niveles de interacción [5].

La efectividad de los entornos educativos digitales es abordada por algunos autores en dimensiones conductual, emocional y cognitiva [6]. La participación en línea puede evaluarse desde dichas dimensiones, ya sea la interacción con el contenido, la cantidad y calidad de las respuestas o el nivel de compromiso con la propuesta, entre otras. Estas también dependen de los factores emocionales que influyen en el desempeño académico [7].

Procesamiento de lenguaje natural

La IA agrupa un conjunto de técnicas orientadas a dotar a los sistemas de la capacidad de realizar tareas que tradicionalmente

requieren inteligencia humana. Dentro de la IA, el machine learning (ML) constituye un enfoque que permite a los modelos aprender patrones a partir de datos sin necesidad de reglas explícitas programadas. Entre las técnicas de ML se incluyen los algoritmos supervisados, no supervisados y de aprendizaje por refuerzo. A su vez, el procesamiento de lenguaje natural (NLP) es una rama de la IA que se enfoca en la interpretación y generación automática de texto o voz en lenguaje humano. El NLP emplea modelos estadísticos y de aprendizaje profundo para tareas como clasificación de textos, análisis de sentimientos o reconocimiento de entidades. En esta categoría emergen los modelos de lenguaje de gran escala (LLM) que utilizan arquitecturas de redes neuronales tipo Transformer, empleadas en este trabajo, y entrenadas con enormes volúmenes de datos. Esta categorización facilita seleccionar la técnica más adecuada según los objetivos y el tipo de datos disponibles. Se muestra a continuación las diferentes categorías de IA.



- Artificial Intelligence (AI)
- Machine Learning (ML)
- Deep Learning (DL)
- Natural Language Processing (NLP)
- Large Language Model (LLM)¹
- Conversational AI (Conv. AI)²

¹LLM is an intersection of DL and NLP

²Conversational AI is a combination of ML and NLP. It may include DL and LLM.

Figura 1: Categorías de inteligencia artificial

Fuente: Gennadiy Bezko, blog MiaRec

El NLP se enfoca en la interacción entre las computadoras y el lenguaje humano, permitiendo a las máquinas interpretar texto, el ML se centra en algoritmos que permiten aprender de datos. Ambos utilizan modelos matemáticos para emular aspectos cognitivos [8] y ser capaces de analizar datos para extraer patrones [9] y tienen la capacidad de procesar textos, identificar emociones, evaluar la calidad de las respuestas [10].

El NLP se aplica para examinar las respuestas textuales de los estudiantes, analizar las transcripciones de las clases grabadas, y evaluar las interacciones de los estudiantes [11]. Facilitar la identificación de emociones es útil para medir el compromiso y el enfoque de los estudiantes [12]. Las universidades han adoptado plataformas digitales que permitan visualizar el avance, la interacción y el compromiso del alumnado [13]. El uso de modelos puede revelar tendencias que serían difíciles de identificar [14]. Los modelos de Redes Neuronales Recurrentes y Transformers (como BERT) son capaces de captar relaciones complejas y patrones de secuencia, permiten una evaluación de las interacciones [15].

LLM (Modelo de lenguaje de gran escala)

Los Modelos de Lenguaje de Gran Escala (LLM, por sus siglas en inglés) constituyen una evolución reciente en el procesamiento de lenguaje natural, capaces de capturar patrones semánticos complejos a partir de su entrenamiento sobre grandes volúmenes de datos textuales. Estos modelos han demostrado buen rendimiento superior en tareas como clasificación, análisis de sentimientos, extracción de entidades y generación de texto coherente.

En el marco de este trabajo, los LLM resultan fundamentales para la construcción de indicadores de desempeño estudiantil, particularmente en la medición de la dimensión emocional. Entre ellos, BERT (Bi-directional Encoder Representations from Transformers), desarrollado por Google AI, constituye una de las herramientas más adecuadas para el cálculo del KPI de Experiencia, ya que permite realizar análisis de sentimientos con una comprensión profunda del contexto.

El análisis de sentimientos aplicado a las interacciones estudiantiles posibilita evaluar la percepción subjetiva de la experiencia educativa. A diferencia de modelos tradicionales, BERT utiliza la arquitectura Transformer para interpretar las palabras considerando simultáneamente su contexto anterior y posterior, lo que lo hace especialmente robusto frente a ambigüedades y expresiones complejas propias del discurso en un entorno áulico. Esto se traduce en una mejor clasificación de emociones positivas, negativas y neutras en frases transcritas durante las clases sincrónicas.

Casos de implementación han mostrado la eficacia de BERT en entornos educativos y productivos. aplicaron BERT al análisis de foros académicos en línea, logrando mejoras en la detección de actitudes emocionales frente a enfoques basados en redes neuronales recurrentes. En el ámbito industrial, BERT ha sido adoptado en sistemas de atención al cliente para clasificar emociones y medir la experiencia del usuario (UX), lo cual evidencia su potencial en el cálculo de métricas emocionales y de satisfacción.

En este trabajo, la elección de BERT se justifica a partir de:

1. Precisión contextual: al ser un modelo bidireccional, capta mejor los matices emocionales en las intervenciones de los estudiantes.

2. Disponibilidad de implementaciones preentrenadas en librerías como Transformers (Google y Hugging Face), lo que facilita su adaptación al dominio educativo.

3. Validación en investigaciones previas, donde se ha demostrado su eficacia en entornos de aprendizaje y análisis de datos textuales.

De este modo, la integración de BERT para el cálculo del KPI de Experiencia (KPI_E) en el gemelo digital áulico permite medir con mayor objetividad la dimensión emocional del estudiante durante la clase, generando información crítica para la retroalimentación en tiempo real tanto a docentes como a instituciones educativas.

Problemática por resolver

En los entornos de educación virtual sincrónica, persiste una limitada capacidad para monitorear y comprender en profundidad las dinámicas de interacción y participación estudiantil en tiempo real. Esta limitación dificulta la toma de decisiones pedagógicas oportunas y basadas en evidencia. Aunque la analítica del aprendizaje ha avanzado en la recolección y análisis de datos educativos, aún existe una brecha significativa en la traducción de estos datos en modelos representacionales que reflejen con fidelidad la complejidad del aula virtual. En este contexto, se plantea como problemática la ausencia de herramientas que permitan modelar dinámicamente el comportamiento colectivo del aula a partir de los datos generados durante las clases.

Según los relevamientos publicados por McMurtrie, la IA se está utilizando con más frecuencia en la educación, especialmente en las instituciones de educación superior que buscan hacer que los cursos con un alumnado numeroso sean más interactivos [16].

A partir de las plataformas de educación en línea se generan grandes volúmenes de datos derivados de las interacciones como foros, chats, y transcripciones de clases grabadas. La disponibilidad de datos para analizarlos objetivamente genera las siguientes preguntas:

- ¿Cuáles son los variables más relevantes para realizar un análisis del desempeño de los estudiantes en ambientes virtuales de aprendizaje?

- ¿Cómo construir indicadores resultan más pertinentes para medir el desempeño de los alumnos?

- ¿De qué manera se generan generar herramientas que permitan construir el modelo digital mediante la utilización de tecnologías de NLP y ML?

- ¿Cómo puede aportar el modelo digital a la comprensión, evaluación o incluso predicción del comportamiento educativo dentro de ese entorno?

Justificación del proyecto

El creciente protagonismo de la educación virtual sincrónica ha planteado nuevos desafíos en cuanto al seguimiento, evaluación y mejora de los procesos educativos. A diferencia de los entornos presenciales, las aulas virtuales carecen de señales observables directas que permitan a los docentes interpretar con precisión el nivel de participación, atención y comprensión de los

estudiantes. Si bien la analítica del aprendizaje ha contribuido significativamente a la recopilación de datos en contextos digitales, persiste la necesidad de traducir dicha información en modelos comprensibles, dinámicos y accionables.

En este sentido, el concepto de gemelo digital —proveniente del ámbito industrial y recientemente introducido en la educación— ofrece un marco prometedor para la representación y monitoreo de los procesos educativos en tiempo real. Este proyecto se justifica por su potencial para integrar enfoques tecnológicos y pedagógicos en una herramienta innovadora, capaz de proporcionar una visión sistémica del aula virtual, optimizar la intervención docente y enriquecer la investigación en el campo de la tecnología educativa y la analítica del aprendizaje.

Objetivos del trabajo

Objetivo General

El objetivo del presente trabajo es el desarrollo de una prueba de concepto de un gemelo digital áulico, construido a partir de datos derivados de la analítica de clases virtuales sincrónicas, con el propósito de modelar dinámicamente las interacciones, comportamientos y patrones del proceso de enseñanza-aprendizaje en entornos digitales.

Esto requiere la elaboración de sistema basado en técnicas de ML y NLP para analizar, clasificar y medir la participación de estudiantes en las clases utilizando como base los registros correspondientes a la grabación de clases sincrónicas en entornos virtuales.

Objetivos secundarios

1. Identificar y categorizar las variables e indicadores relevantes generados en clases virtuales sincrónicas que permitan caracterizar el comportamiento del aula.
2. Diseñar un modelo representacional del aula virtual (gemelo digital áulico) utilizando técnicas de analítica del aprendizaje y herramientas de visualización dinámica.
3. Evaluar la utilidad y aplicabilidad del gemelo digital áulico como herramienta de apoyo para la toma de decisiones pedagógicas en tiempo real.

Desarrollo del trabajo

El trabajo se estructura en cinco etapas ejecutadas secuencialmente. Se utiliza Python como lenguaje de programación por considerarse adecuado por sus características. Se trata de un lenguaje interpretado, de alto nivel, de tipado dinámico, multiplataforma, de código abierto, orientado a objeto, gran comunidad, diversidad librerías, etc [15].

Recolección de Datos

Se trabajó sobre un conjunto de datos no estructurados provenientes de los textos resultantes de las transcripciones la recopilación de las diez clases grabadas de física correspondientes al primer año de ingeniería de la UNLAM del curso de verano.

Como herramienta de grabación de los registros de audio se aprovechó la plataforma de colaboración MS Teams, que es la misma que regularmente utiliza la universidad para dictar las clases sincrónicas virtuales. Esta genera los registros de las grabaciones en formato de video mp4. A partir de allí se extrajo el audio para su posterior procesamiento en formato mp3.

El primer paso consistió en transcribir las automáticamente para construir el conjunto de datos a analizar, proceso conocido como STT por sus siglas en inglés (speech to text). Para ello se utilizó una integración con la librería www.assemblyai.com mediante API pública.

```
import assemblyai as aai
import pandas as pd
import nltk
from nltk.tokenize import sent_tokenize
#Definir llave de API
aai.settings.api_key = "87808a54e93945f2b2632cb01e740bd7"
#Seteo idioma y activo la detección de hablantes. Agrego rutas de files
config = aai.TranscriptionConfig(language_code="es", speaker_labels=True)
audiofile_path1=r"C:\Users\rosal\Downloads\Clase1_Fisica1_prueba.mp3"
audiofile_path2=r"C:\Users\rosal\Downloads\Clase2_Fisica1_prueba.mp3"
audiofile_path3=r"C:\Users\rosal\Downloads\Clase3_Fisica1_prueba.mp3"
audiofile_path5=r"C:\Users\rosal\Downloads\clase5_Fisica1_prueba.mp3"
audiofile_path6=r"C:\Users\rosal\Downloads\clase6_Fisica1_prueba.mp3"
audiofile_path9=r"C:\Users\rosal\Downloads\clase9_Fisica1_prueba.mp3"
audiofile_path10=r"C:\Users\rosal\Downloads\clase10_Fisica1_prueba.mp3"
#Listas con las que vamos a trabajar
listaPreguntas=[]
listaRespuestas=[]
listaHablantes=[]
listaDiscursos=[]
#Configuraciones iniciales
def generarTranscripcion(audiofile_path,config,nombreArchiv):
    transcriber = aai.Transcriber(config=config)
    transcript = transcriber.transcribe(audiofile_path)
    transcripcion_.txt=transcript.text #Genera texto(string)
#Almacenamos los textos y los hablantes en listas
for utterance in transcript.utterances:
    speaker = utterance.speaker
    text = utterance.text
    listaHablantes.append(speaker)
    listaDiscursos.append(text)
```

Figura 2: Parte del script de integración API

Fuente: Elaboración propia

Preparación de Datos

La preparación de los datos hace referencia a un primer procesamiento. En esta etapa es donde se realizan las transformaciones necesarias para adecuar el texto al análisis posterior, esto incluye los siguientes procesos:

Limpieza de texto: elimina errores, caracteres especiales, no deseados, y otros que afecten.

Tokenización: segmenta el texto en unidades más pequeñas palabras, frases o párrafos.

Eliminación de stopwords: elimina palabras vacías que no aportan información relevante.

Lematización/Stemming: transforma cada palabra a su forma raíz, para que identificar las variaciones de una misma palabra.

Identificación de turnos del habla: segmenta las intervenciones identificando al hablante.

Para ilustrar algunos resultados se muestra en el siguiente gráfico algunas intervenciones de estudiantes y docente discriminadas por hablante. Es preciso aclarar que el hablante A se refiere al docente, mientras que el hablante B y C son estudiantes.

	Hablante	Preguntas
0	A	¿Cuál es el valor de la aceleración normal de ...
1	A	¿Cuál es la velocidad de la pelota al impactar... ¿Esto está visto desde arriba, no?
2	A	Si yo corto la cuerda acá, para dónde se va l...
3	A	¿Estoy entendiendo bien el enunciado?
4	A	...
...
2195	A	¿Está bien?
2196	A	¿Alguna duda, alguna inquietud, algo que quieran?
2197	A	¿Quieren parar 5 min o siguen derecho?
2198	A	¿Cómo hacerlo normalmente?
2199	A	¿Cuánto tiempo ponemos?
[2200 rows x 2 columns]		
	Hablante	Discurso
0	A	Ya el dibujito medio que me asusta. Una pelota...
1	B	¿Tipo en diagonal con la tangencial?
2	A	¿Para allá, no? Está bien. Entonces, esta velo...
3	B	No lo estaba entendiendo porque no lo veía des...
4	A	Lo leo de nuevo por las dudas. La pelota acá, ...
...
1315	A	Ahí en algún otro, en alguna otra instancia.
1316	C	Buenísimo, profe, muchas gracias.
1317	A	Muchas gracias.

Figura 3: Muestra de intervenciones de estudiantes y del docente

Fuente: Elaboración propia

Construcción de indicadores

Esta tercera instancia refiere a la elaboración de los indicadores a partir de las variables referenciadas. Es importante aclarar que suele hacerse referencia a los indicadores de desempeño como KPI por sus siglas en inglés (Key Performance Indicator).

A lo largo del trabajo se consideran tres aspectos clave del estudiante conductual, emocional y cognitiva. A partir de ello se definen tres indicadores clave.

1. Compromiso: indicador referido al aspecto conductual. Se medirá a través de la cantidad de interacciones verbales, ya sea del docente como de estudiantes generando y respondiendo preguntas.

2. Experiencia: indicador referido al aspecto emocional. Se determinará a partir del análisis de la intensidad emocional utilizando técnicas de análisis de **sentimiento**.

3. Enfoque: indicador referido al aspecto cognitivo. Se evaluará mediante la relevancia de las intervenciones de los estudiantes en relación con las unidades temáticas tratadas.

Indicador de compromiso

Se trata de un indicador conductual, que refiere al grado de involucramiento de los estudiantes con las actividades propuestas. Los modelos de NLP pueden medir el compromiso mediante el análisis de la frecuencia de interacción y la pertinencia de las intervenciones dentro de la clase [17].

Se contabiliza la cantidad de preguntas detectadas, ya sean formuladas y/o respondidas, y se obtiene un nivel de compromiso de cada alumno. Dado que cada una de estas métricas tiene diferentes escalas, es preciso normalizarlas para generar un indicador consolidado a partir de ellas.

Se mide la cantidad de interacciones [18], se utiliza la normalización Min-Max [19] y luego se calcula el valor del indicador de compromiso.

Se define al indicador conductual como KPI_compromiso

Este KPI de compromiso evalúa el nivel de participación activa del estudiante durante las clases.

Se define KPI de Compromiso = KPI_C

$$KPI_C = (N_{int} + N_{preg} + N_{resp}) / T_{tot}$$

- N_int: Número de intervenciones del estudiante.
- N_preg: Número de preguntas formuladas.
- N_resp: Número de respuestas dadas.
- T_tot: Total de turnos de habla registrados en clase.

Rango estimado de interpretación

- < 0.30: Participación baja
- 0.31 – 0.60: Participación moderada
- > 0.60: Participación alta

Se observa en la siguiente gráfica el indicador compromiso de los estudiantes con mayor participación en las diferentes clases.

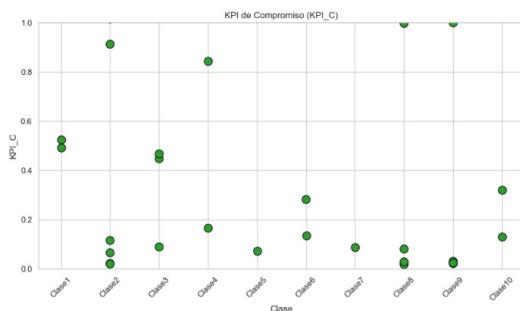


Figura 4: Distribución del índice KPI_C

Fuente: Elaboración propia

Indicador de experiencia

Se trata de un indicador emocional que refiere a la percepción de la experiencia subjetiva que los estudiantes tienen en la clase. El análisis de los comentarios y de sus intervenciones proporciona información sobre su nivel de satisfacción. Las técnicas de topic modeling (modelo de temas) permite identificar los temas de interés [20]. En este caso el algoritmo utilizado es el modelo de análisis de sentimientos BERT (Bidirectional Encoder Representations from Transformers) que utiliza dispone de capacidad de comprender el contexto bidireccional de las palabras en una secuencia.

```
#Realizamos análisis de sentimientos de estos dos dataframes
from transformers import pipeline

# Cargamos el modelo preentrenado de análisis de sentimientos
classifier = pipeline('sentiment-analysis',
                      model="nlptown/bert-base-multilingual-uncased-sentiment",
                      truncation=True) # Truncar textos largos a 512 tokens

def obtener_polaridad_score(pregunta):
    result = classifier(pregunta) # Realiza el análisis de sentimiento
    # Devolvemos el label (polaridad) de la respuesta, puede ser 'LABEL_0', 'LABEL_1', etc.
```

Figura 5: Muestra de aplicación de uno de los KPI

Fuente: Elaboración propia

Se define al indicador emocional como KPI_experiencia.

En la siguiente figura se ilustra el resultado de las clasificaciones con algunos ejemplos.

Filtrado5_estrellas.png					
Hablante		Discurso	Valoración	Score	
30	B	Buenísimo, profe, gracias.	5 stars	0.842238	
32	B	Sí, sí, buenísimo.	5 stars	0.759194	
63	B	Buenísimo, profe.	5 stars	0.771217	
140	B	Si. Muchísimas gracias, profeta.	5 stars	0.604452	
211	A	Exactamente.	5 stars	0.650418	
239	E	Buenísimo.	5 stars	0.761205	
340	A	Perfecto. Y está lleno de situaciones. ¿Cuando...	5 stars	0.706222	
427	A	Perfecto. A ver, bueno, denme 1 s y le voy a s...	5 stars	0.794192	
430	C	Buenísimo, entonces. Muy amable, profe, muchis...	5 stars	0.601502	
432	C	Perfecto, muy amable, Gustavo.	5 stars	0.891618	
524	E	Buenísimo, buenísimo.	5 stars	0.754680	
533	A	Gracias. Bueno, muy bien, perfecto. ¡Nadie más...	5 stars	0.779368	
564	B	Buenísimo, profe, buenísimo.	5 stars	0.807030	
583	D	Listo, perfecto.	5 stars	0.778544	
604	E	Dale, buenísimo, profe.	5 stars	0.699701	
685	C	Genial, profe, genial. La parte de la trigonom...	5 stars	0.715561	
751	C	Perfecto, gracias.	5 stars	0.872111	
765	C	Sí, buenísimo, gracias.	5 stars	0.723557	

Figura 6: Muestra de respuestas identificadas y su valoración

Fuente: Elaboración propia

Este indicador pretende medir la calidad emocional en las interacciones del estudiante durante la clase.

Se define KPI de Experiencia = KPI_E

$$KPI_E = ((S_{\text{pos}} - S_{\text{neg}}) / S_{\text{tot}})$$

• S_{pos} : Sentimientos positivos detectados.

• S_{neg} : Sentimientos negativos detectados.

• S_{tot} : Total de frases analizadas en lo emocional

Rango estimado de interpretación

- < 0.5 : Experiencia deficiente
- $0.51 - 1$: Experiencia aceptable
- > 1 : Experiencia enriquecida

Se observa en la siguiente gráfica los KPI_E de los estudiantes con los valores más altos en las diferentes clases.

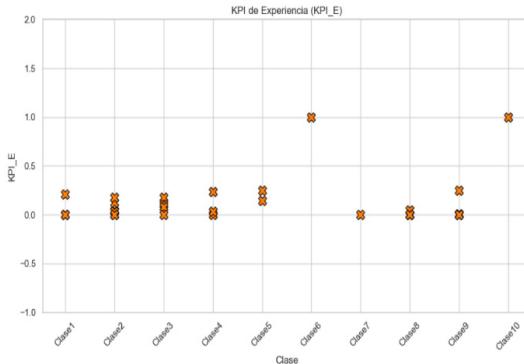


Figura 7: Distribución del índice KPI_E

Fuente: Elaboración propia

Indicador de enfoque

Se trata de un indicador cognitivo que pretende medir la atención que los estudiantes mantienen sobre el contenido de la clase. A través de las interacciones permite inferir si un estudiante está siguiendo la temática del curso. Se construye mediante la relevancia de intervenciones de los estudiantes en relación con unidades temáticas [21]. Se clasifican los tipos de preguntas para luego relacionarlas con temáticas particulares.

Se realizó el modelado de ML mediante la librería scikitlearn de Python con algoritmo de Naives Bayes para clasificar los tipos de preguntas. Este indicador de enfoque pretende medir la alineación temática de las intervenciones con los objetivos curriculares de la clase.

```
# Iterar sobre las preguntas en el DataFrame
for i, pregunta in enumerate(df_Preguntas["Preguntas"]):
    doc = nlp(pregunta)
    texto_encontrado = None
    tematica_encontrada = None

    # Iterar sobre las entidades reconocidas
    for ent in doc.ents:
        # Verificar si la entidad coincide con las claves del diccionario
        if ent.text in tematicas_dict:
            texto_encontrado = ent.text
            tematica_encontrada = tematicas_dict[ent.text]
            break # Detener la búsqueda al encontrar la primera coincidencia

    # Asignar los valores encontrados a la fila actual del DataFrame
    df_Preguntas.loc[i, "Texto"] = texto_encontrado
    df_Preguntas.loc[i, "Tematica"] = tematica_encontrada

# Mostrar el DataFrame actualizado
df_preg_cinemática=df_Preguntas[df_Preguntas['Tematica'] == 'Cinemática']
df_preg_dinámica=df_Preguntas[df_Preguntas['Tematica'] == 'Dinámica']

#Cantidad de preguntas realizadas de cada temática
cantpreg_x_tematica=df_Preguntas.groupby('Tematica').size()
print(cantpreg_x_tematica)

display(df_preg_cinemática)
display(df_preg_dinámica)
```

Figura 8: Segmento del script para identificar unidades temáticas

Fuente: Elaboración propia

Se define KPI de Enfoque = KPI_F

$$KPI_F = (C_{clave} / C_{tot}) * R_{coh}$$

- C_{clave} : Número de conceptos clave mencionados.
- C_{tot} : Total de conceptos tratados en clase.

- R_{coh} : Índice de coherencia semántica entre segmentos del discurso (calculado comparado con diccionario de términos).

Estas fórmulas pueden ser ajustadas/ ponderadas según el curso, estilo docente o criterios institucionales.

Rango estimado de interpretación:

- < 0.40 : Bajo enfoque
- $0.41 - 0.70$: Enfoque intermedio
- > 0.70 : Enfoque elevado

Se observa en la siguiente gráfica los KPI_F de los estudiantes con los valores más altos en las diferentes clases.

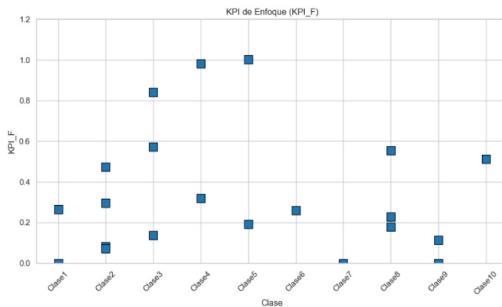


Figura 9: Distribución del índice KPI_F

Fuente: Elaboración propia

Evaluación y validación

Para garantizar la precisión y fiabilidad de los modelos, se emplea validación cruzada (cross-validation), permite dividir los datos en subconjuntos de entrenamiento y prueba.

Para la evaluación y validación del modelo de datos se procedió a dividir el conjunto de datos original en dos subconjuntos independientes. El 80% de los registros se destinó al entrenamiento del modelo, permitiendo el ajuste de sus parámetros internos. El 20% restante se reservó para la fase de prueba, garantizando una evaluación imparcial de su capacidad de generalización. Adicionalmente, se aplicó una validación cruzada k-fold para reducir la varianza asociada a una única partición de datos. Este procedimiento permitió medir el desempeño promedio del modelo y su estabilidad frente a cambios en el conjunto de entrenamiento. La validación también contempló el análisis de la matriz de confusión para identificar patrones de error. Con ello se aseguró que el modelo no presentara sesgos significativos.

En el siguiente gráfico se observa la matriz de confusión. Es una representación visual del resultado de la clasificación del

modelo en cuanto a los tipos de preguntas formuladas por los estudiantes. Las filas representan las clases reales y las columnas las clases predichas por el modelo.

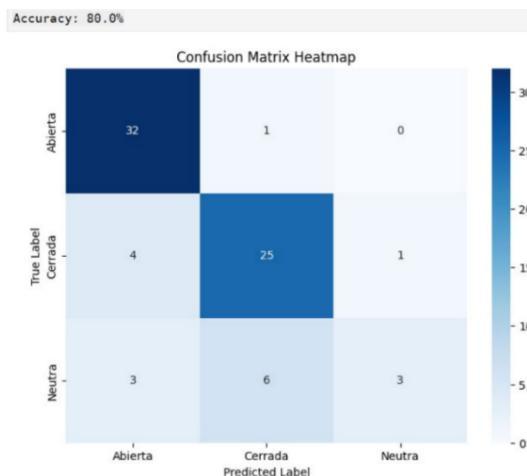


Figura 10: Matriz de confusión

Fuente: Elaboración propia

Propuesta para la construcción del Gemelo Digital Áulico

El **gemelo digital** es una representación virtual dinámica de un sistema real, que se actualiza de manera continua con datos generados en tiempo real. En este caso, el sistema real está constituido por el estudiante y su interacción en clases virtuales sincrónicas. El gemelo digital permite modelar, simular y predecir la evolución del desempeño académico de forma individual o grupal, integrando información conductual, emocional y cognitiva.

En este trabajo, se propone un **gemelo digital del estudiante** que se construye a partir de los indicadores previamente definidos:

- **KPI_C (Compromiso – conductual)**: mide el nivel de participación activa en clase.
- **KPI_E (Experiencia – emocional)**: refleja la calidad emocional de las interacciones.
- **KPI_F (Enfoque – cognitivo)**: evalúa la alineación temática y nivel de atención.

A medida que se registran nuevas interacciones (transcripciones de clases, turnos de habla, preguntas y respuestas), el gemelo digital se actualiza en tiempo real, generando un **perfil dinámico y evolutivo** de cada estudiante.

Beneficios del Gemelo Digital en Educación

1. **Monitoreo personalizado**: permite a los docentes contar con un “espejo digital” del estudiante para identificar patrones de aprendizaje y áreas de mejora.
2. **Detección temprana de riesgos**: el sistema alerta cuando un estudiante presenta baja participación, emociones negativas persistentes o falta de enfoque.
3. **Simulación de escenarios**: posibilita predecir cómo cambiaría el desempeño bajo diferentes condiciones pedagógicas (más interacción, variación en contenidos, uso de recursos adicionales).
4. **Optimización de la enseñanza**: los docentes reciben retroalimentación basada en datos objetivos para ajustar en tiempo real la estrategia de enseñanza.

5. **Escalabilidad institucional**: el modelo puede integrarse a sistemas LMS, facilitando el seguimiento de cohortes completas de estudiantes en entornos virtuales.

El flujo de funcionamiento del modelo puede representarse en el siguiente esquema:

Estudiante (interacción en clase)



Recolección de datos (audio, texto, transcripciones)



Procesamiento NLP + ML

(tokenización, análisis de sentimiento, clasificación)



Cálculo de indicadores (KPI_C, KPI_E, KPI_F)



Gemelo Digital del Estudiante

(perfil dinámico, actualizado en tiempo real)



Retroalimentación

- Estudiante: recomendaciones personalizadas
- Docente: alertas e insights sobre la clase
- Institución: reportes agregados

Los datos conductuales, emocionales y cognitivos de los estudiantes son procesados mediante técnicas de NLP y ML para generar indicadores clave (KPI de Compromiso, Experiencia y Enfoque). Estos alimentan la construcción del Gemelo Digital del Estudiante, que se actualiza de forma dinámi-

ca y brinda retroalimentación diferenciada a docentes, estudiantes e instituciones educativas.

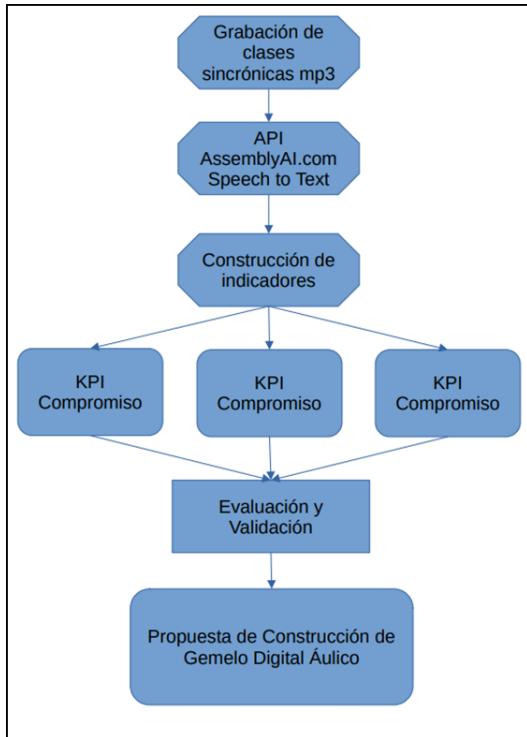


Figura 11: Estructura del desarrollo del trabajo

Fuente: Elaboración propia

Resultados

En la distribución de la variable preguntas formuladas por clase se observa una clara preponderancia del hablante A, identificado como el docente. El resto de los hablantes, desde B, C, etc en adelante, corresponden a diferentes alumnos.

La participación de los estudiantes en términos de preguntas formuladas alcanza su máximo en la clase 9, precisamente la que corresponde como clase previa a la fecha prevista de examen.



Figura 12: Distribución de preguntas realizadas

Fuente: Elaboración propia

En el siguiente gráfico se observa la distribución de respuestas y se destaca cierto balance entre las respuestas del docente y la de los alumnos en la mayoría de las clases. En la clase 9 se destaca la participación del alumno identificado como hablante B siendo su cantidad de intervenciones aún mayor a la del docente (hablante A).

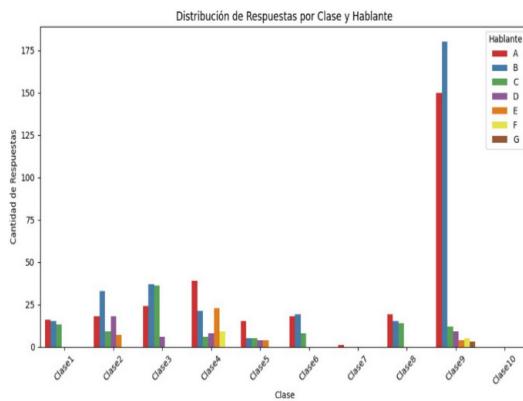


Figura 13: Distribución de respuestas formuladas

Fuente: Elaboración propia

Para enriquecer el análisis hacemos resulta interesante realizar un cruce entre preguntas y las respuestas de docente y estudiante respectivamente.

Para eso se registra en la siguiente figura y se comprueba que no existe adecuado balance entre ambas. Es importante aclarar que se grafica la cantidad de respuestas de alumnos en función de preguntas de docente para las distintas clases. Se observa que la cantidad de preguntas del docente superan en un orden de magnitud a las respuestas de los estudiantes.

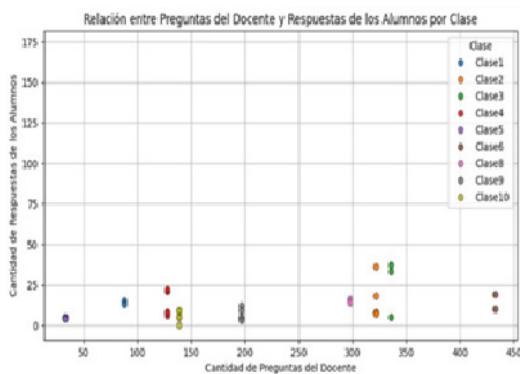


Figura 14: Cruce de preguntas de docente y respuestas de alumnos

Fuente: Elaboración propia

Se realiza ahora un segundo cruce de variables, de forma de graficar la cantidad de preguntas formuladas por los estudiantes en función de las respuestas del docente. En este caso se asemeja a recta con pendiente de $m=0.5$, lo que refleja cierto equilibrio en la cantidad de participaciones, por cada pregunta de alumno hay dos respuestas del docente.

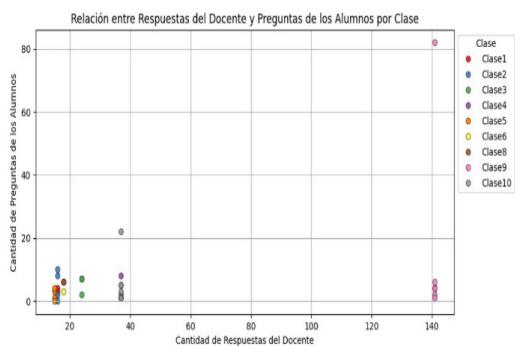


Figura 15: Cruce de preguntas de alumnos y respuestas de docente

Fuente: Elaboración propia

Conclusiones

La integración de analítica del aprendizaje con modelos de representación digital, como el gemelo digital áulico, ofrece una vía innovadora para abordar los desafíos de monitoreo y evaluación en entornos de educación virtual sincrónica. Esta propuesta demuestra la pertinencia de desarrollar representaciones dinámicas que traduzcan los datos generados en las plataformas educativas en información útil para la toma de decisiones pedagógicas.

El uso combinado de herramientas como NLTK y BERT permite abordar distintas dimensiones del proceso educativo, en particular la experiencia subjetiva del estudiante y su nivel de enfoque durante las sesiones virtuales. Mientras que NLTK facilita un análisis léxico y sintáctico básico orientado a indicadores de satisfacción, BERT posibilita un análisis semántico profundo que enriquece la comprensión del comportamiento y la atención estudiantil.

La construcción de un gemelo digital áulico se configura como una herramienta estratégica para el acompañamiento en tiempo real del proceso de enseñanza-apren-

dizaje. Su capacidad para modelar las interacciones y patrones del aula virtual puede contribuir a una retroalimentación inmediata, una gestión más eficiente del tiempo docente y una personalización del aprendizaje más informada.

Desde el punto de vista metodológico, esta prueba de concepto propone un enfoque transdisciplinario, articulando elementos de la inteligencia artificial, la analítica de datos educativos y la teoría del aprendizaje, lo que refuerza su valor tanto científico como práctico en el ámbito de la tecnología educativa.

Finalmente, esta propuesta abre nuevas líneas de investigación en torno al diseño de ambientes de aprendizaje mediados por tecnologías inteligentes, donde la representación computacional del aula no solo describe, sino que también predice y optimiza las condiciones para el aprendizaje significativo en contextos virtuales.

Los resultados obtenidos han validado que es posible clasificar los niveles de participación de los estudiantes y docentes utilizando las tecnologías referenciadas y los indicadores propuestos tomados como referencia.

El análisis de los indicadores de compromiso, experiencia y enfoque permitió obtener una visión integral del desempeño de los estudiantes en entornos de clase sincrónica virtual. Los resultados evidenciaron que un mayor nivel de compromiso se asocia con una participación más activa y sostenida a lo largo de la sesión. El indicador de experiencia reflejó la percepción de los estudiantes sobre la claridad, interactividad y utilidad de las actividades propuestas. Por su parte, el indicador de enfoque mostró correlación directa con la capacidad de atención

y la relevancia de las intervenciones con los contenidos curriculares. En conjunto, estos tres indicadores ofrecieron una base sólida para identificar patrones de participación y áreas de mejora pedagógica. Se observó que la interacción bidireccional y el uso de recursos multimedia favorecen tanto la experiencia como el enfoque. Además, la comparación longitudinal de sesiones permitió detectar oportunidades de mejora.

Además, el trabajo ofreció una aproximación al contexto productivo de las técnicas del NLP con la limitación que depende de un posprocesamiento, si bien resulta útil para una evaluación no aplicada en tiempo real.

La combinación de NLP para el análisis de texto y ML para la clasificación de las interacciones ha permitido identificar patrones relevantes en el desempeño de los estudiantes su compromiso en la participación activa, como la intensidad emocional en las respuestas, y la calidad de la experiencia educativa.

La metodología propuesta no solo permite identificar patrones de participación y áreas de mejora, sino que también aporta un marco replicable en diferentes contextos educativos. De allí que está previsto continuar con la recopilación de datos para seguir nutriendo al corpus con las transcripciones de las clases correspondientes al primer cuatrimestre de la cursada para seguir entrenando al modelo y mejorar su comportamiento.

Como futuros trabajos está previsto avanzar en el análisis predictivo de las tasas de graduación, las tasas de abandono y el rendimiento académico. Se considera la posibilidad de construir un gemelo digital áulico para conocer las potenciales respuestas de los estudiantes frente a diversas

situaciones educativas. También se propone un sistema que además permitiría la intervención en tiempo real de lo que está sucediendo en clase como ayuda en línea al docente. última sección del artículo deben ser las Conclusiones. Puede incluir las líneas de trabajo/investigación a seguir.

Referencias

- [1] Andresen S.L. (2002) John McCarthy: father of Artificial Intelligence. *IEEE Intelligent Systems* (vol 17, art 5, pag 84 - 85)
- [2] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A., Kaiser, Ł., & Polosukhin, I. (2017). "Attention is all you need." *Proceedings of the 31st International Conference on Neural Information Processing Systems* (NeurIPS 2017), 5998-6008.
- [3] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding." *Proceedings of NAACL-HLT*, 4171-4186
- [4] Chen, L., Chen, P., & Lin, Z. (2020). Artificial Intelligence in Education: a review. *IEEE Access*, 8, 75264-75278.
- [5] Martin, F., Sunley, R., & Turner, P. (2019). "Student engagement in online learning environments: A literature review." *The International Journal of Information and Learning Technology*, 36(4), 284-29.
- [6] Fredricks, Blumenfeld & Paris, A. H. (2004). "School engagement: Potential of the concept, state of the evidence." *Review of Educational Research*, 74(1), 59-109
- [7] Bower, M., Cavanagh, M., & Moloney, R. (2015). "Designing a Blended Learning Environment: A Case Study of Teaching Business Information Systems." *International Journal of Educational Technology in Higher Education*, 12(1), 1-15
- [8] Woolf, B. P. (2010). *Building Intelligent Interactive Tutors: Student-Centered Strategies for Learning*. Ed. Elsevier.
- [9] Davenport, T. H., & Ronanki, R. (2018). "Artificial intelligence for the real world." *Harvard Business Review*, 96(1), 108-116
- [10] Jurafsky, D., & Martin, J. H. (2020). *"Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition"*. Ed. Pearson.
- [11] Baker, R. S., D'Mello, S. K., Rodrigo, M. M., & Graesser, A. C. (2019). "Better understanding student engagement and learning through data mining and machine learning techniques." *Computers in Human Behavior*, 100, 68-75.
- [12] Machova K., Szaboova M., Paralic J., & Mikco (2023). "Detection of emotion by text analysis using machine learning." *Page 3*.
- [13] Siemens, G. (2013). "Learning analytics: The emergence of a discipline." *American Behavioral Scientist*, 57(10), 1380-1400
- [14] McMurtrie Beth (2020). Preparing for different kind of classroom. The Chronicle of higher education Inc. (art 1, pag 5-8). Washington
- [15] Baker, R.S., & Yacef, K (2009). "The state of educational data mining in 2009". Proceedings of 2nd Conference on educational Data Mining, 3-13
- [16] Bel W., *Algoritmos y estructuras de datos en Python*. Un enfoque ágil. Facultad Ciencia y Tec. (2020). pag17. Ed Uader.
- [17] Garrison, D. R., Anderson, T., & Archer, W. (2000). "Critical thinking, cognitive presence, and computer conferencing in distance education." *American Journal of Distance Education*, 15, 7-23

[18] Freeman, S., et al. (2014). *Active learning increases student performance in science, engineering, and mathematics*. Proceedings of the National Academy of Sciences, 111(23), 8410-8415.

[19] Kotsiantis, S. B., & Pintelas, P. E. (2004). *Data Preprocessing for Supervised Learning*. International Journal of Computer Science, 1(1), 111-116.

[20] Zhao X., Jiang Z. & Gray J., (2020) “*Text Classification and Topic Modeling for Online Discussion Forums*”. Trends and Applications of Text Summarization Techniques. (ch 6) 151-166

[21] Blei, D. M., Ng, A. Y., & Lafferty, J. D. (2003). “Latent Dirichlet allocation.” *Journal of Machine Learning Research*, 3, 993-1022