

The image shows a modern classroom with students seated at desks, each with a laptop or tablet. They are looking towards a large interactive screen at the front of the room. The screen displays a lesson titled "BRIDGE THE GAP" with a diagram of a bridge and a pie chart. The screen also shows a sidebar with various educational tools like "Question of the Day", "Calendar", and "Class Resources". The logo for "EDUCA" is visible in the top left corner of the image.

EDUCA

Acción empresarial
por la educación!

Nota de trabajo No. 35:

EDUCACION Y MACHINE LEARNING: La puerta de entrada a un nuevo paradigma

Autor: **Enrique Darwin Caraballo, CEO EDUCA**

Los diferentes instrumentos de medición de los niveles de aprendizaje de los estudiantes que hoy disponemos en América Latina -léase pruebas diagnósticas al finalizar cada ciclo escolar y aplicadas por las autoridades competentes en cada país, las pruebas de calidad de los aprendizajes suministradas a nivel regional por la UNESCO o las pruebas PISA, impulsadas por la OCDE- confluyen en una similar y dramática conclusión: los estudiantes latinoamericanos no están aprendiendo lo suficiente. Estas mediciones se suceden, una tras otra, arribando siempre a esta conclusión, con escasos niveles de variación y sin un patrón constante que permita intuir, aunque lenta, una evolución positiva. Ante esta situación las autoridades educativas hacen, al decir de Eric Hanusheck (1981) lo que mejor saben hacer. Es decir, echar dinero a los problemas. Así, los presupuestos educativos, públicos y privados, en toda la región han experimentado incrementos significativos, sin el consecuente efecto sobre los niveles de aprendizaje de los estudiantes.

En esencia, las dinámicas de los procesos de aprendizaje no han variado desde cuando la educación se expandió al conjunto de la población y dejó de ser el privilegio de los grupos más aventajados de la sociedad. La masificación de la educación se corresponde con la instalación plena de la revolución industrial, con la consolidación de los estados nacionales y la expansión y dominio del estado de derecho. El cometido de la educación era formar trabajadores, ciudadanos integrados a una identidad nacional, conocedores y respetuosos de la ley. En ese modelo, el conocimiento debía experimentar una “transposición didáctica”, desde el maestro al alumno (Yves, Chevallard; 1997).

Es decir, el conocimiento se concebía en poder monopólico del profesor, que actuaba como agente activo y emisor; mientras que el estudiante desempeñaba un rol pasivo, en tanto receptor.

La irrupción de la tecnología y el auge de los medios de comunicación, echan por tierra los pilares de este paradigma. En efecto, los estudiantes hoy disponen de cientos de canales para recibir saberes, contrastarlos y cuestionarlos con los que plantea el docente en el aula. Ante esta anomalía evidente del paradigma vigente, los sistemas han pretendido tratar dicha anomalía como una simple alteración, no significativa, y absorber sus efectos, simplemente entregando dispositivos o agregando mayor tiempo en las aulas.

En paralelo, la ciencia viene aprendiendo y comprendiendo de manera exponencial el funcionamiento del cerebro humano. También, sobre los procesos bioquímicos que intervienen en el aprendizaje, los cuales son tenuemente incorporados a la hora de definir políticas educativas. Igualmente, el proceso de toma de decisiones en este sector, fuertemente influenciado por educadores formados en áreas de la psicología, son renuentes a considerar la data para guiar acciones, y se limita su consideración a un uso instrumental, solo cuando es congruente con la cosmovisión de carácter ideológico que adopta la autoridad o el gerente de turno.

Tal y como lo plantea Fernando Sandoya (MIT), la solución para una mayor efectividad de los procesos de aprendizaje, bien podría yacer en la ciencia de datos y en el diseño de algoritmos de Machine Learning (ML); sea de clasificación o de otra especie, que le permitiera a cada individuo desarrollarse allí donde demuestra mayor talento natural e intereses. Este es el desafío y el elemento diferenciador del paradigma que se avecina. Es decir, romper con la estandarización del perfil de egreso de cada nivel educativo. Paradigma, diseñado como una cadena de producción industrial cuyo resultado al final del proceso debía basarse en la estandarización y patrones de calidad definidos por un único tipo. El ML puede contribuir a personalizar propuestas educativas para que cada alumno oriente su proceso educativo en función de sus intereses, atributos y talentos. El ML puede convertirse así, en uno de los instrumentos más poderosos con que cuente el docente en el aula para guiar a sus estudiantes, y para las autoridades educativas en sus responsabilidades de planeación y ejecución de políticas. En este sentido, el ML puede constituir una de las claves para reinventar la forma en cómo educamos y aprendemos.

Antes de describir e identificar las 4 etapas que definen el proceso de ML, en el sector educación es preciso definir una etapa 0, inicial, la captura de la data. Lo que puede parecer evidente en otras disciplinas, aquí la forma de captura y los instrumentos utilizados son de vital importancia. Aún los sistemas siguen aplicando pruebas de los estudiantes en papel, capturando datos administrativos en el mismo medio, recogidos por personas no entrenadas y no considerando decenas de dimensiones relevantes que tanto la teoría como la evidencia demuestran que impactan en los aprendizajes. Me refiero a los patrones alimenticios y de sueño; a la composición de la familia y características del hogar del estudiante; al nivel de desarrollo psicocerebral; a la evaluación del número y velocidad de las redes neuronales de cada individuo; a los balances bioquímicos, tanto a nivel cerebral como a nivel orgánico general de los alumnos; de igual manera no se relevan intereses, preferencias, y opiniones de los propios estudiantes. Asimismo, tampoco se recaba data de esta naturaleza en los docentes y padres de familia. De modo que en este sector se debe hacer un esfuerzo adicional para repensar los medios, las formas y los atributos que son relevantes recabar para ganar en contenido, validez y confiabilidad de la data relevada.

Una vez resuelto el problema de la captura y calidad de la información recabada, corresponde comprender esos datos. Como se trata de un fenómeno complejo, multidimensional y que deberá administrar cientos de atributos de cada perfil, constituirá un reto formular las preguntas adecuadas, y presentar las diferentes asociaciones o factores que determinan la variable dependiente, ergo el nivel de aprendizaje, de forma amigable y comprensible. Sobre todo, teniendo en cuenta que los tomadores de decisión, en general, no cuentan con la formación para interpretar datos de forma compleja. Aplicar el método de análisis de componentes principales será mandatorio en esta fase del proceso.

Una forma de mejorar la capacidad de análisis y contribuir a formular mejores preguntas a la data, es recurrir al aprendizaje no supervisado. Es decir, definir una herramienta o algoritmo que permita a los sistemas informáticos determinar patrones, relaciones entre los datos y estructuras no evidentes ni consistentes con las hipótesis tradicionales formuladas a partir de la teoría. Una estrategia que apuntala a esa dirección es Learning Analytics, un método para comprender y optimizar el proceso de aprendizaje de los estudiantes a través del procesamiento y visualización de datos sobre las características de los

estudiantes y su progreso y trayectoria educativa. Esta herramienta es principalmente utilizada por docentes para identificar deficiencias en los aprendizajes, e inclusive sugerirles currículos personalizados para sus pupilos. También, es usado frecuentemente por estudiantes para traquear sus propios progresos en término de aprendizaje (Clow, 2012). Sería interesante también estudiar la viabilidad de construir una ontología informática de los datos para a su vez se puedan establecer jerarquías entre los datos a considerar.

Con buena data disponible, correctamente presentada y comprendida estarán dadas las condiciones para predecir el comportamiento de los datos en el futuro. En educación se recurre a los modelos de predicción lineales o regresiones simples, también al análisis factorial. Sin dudas necesario pero insuficiente. Entiendo, hay una enorme oportunidad para aplicar Deep Learning, a partir de la creación de redes neuronales que sean capaces de predecir el desempeño de un estudiante para lograr introducir los correctivos a tiempo. Esto representará sin dudas un desafío logístico para el docente en el aula, como también a nivel sistémico para promover cambios metodológicos a tiempo, de forma tal que el nivel de fracaso estudiantil pronosticado sea minimizado.

Tomar decisiones en función de los modelos predictivos supone en este sector, además de costos administrativos y financieros costos políticos y reputacionales. En ocasiones los tomadores de decisión son renuentes a ensayar cambios radicales o significativos por el grado de incertidumbre que supone iniciar alternativas y los costos asociados. A estos costos debe agregarse aquellos relacionados con la resistencia al cambio de agentes que pueden sentirse afectados por los cambios sugeridos. Sin dudas, en este sector, se deben extremar las medidas y tomar los recaudos necesarios para minimizar los costos diversos y al mismo tiempo intentar explotar los modelos más efectivos para contagiar al conjunto del sistema. Por esta razón la aplicación en pequeñas escalas, y en ambientes lo más controlables posibles suelen ser las estrategias más adecuadas previo a impactar al conjunto de una política pública o sistema.

Experimentar en escalas reducidas y recurriendo a grupos de control constituye la mejor herramienta para entender las inferencias causales en sectores como la educación cuya naturaleza no permite controlar todos los elementos que intervienen, en este caso en los aprendizajes. Por esta razón se

habla de diseños cuasiexperimentales (Stanley y Campbell, 1963) para fenómenos sociales en los cuales son válidos los controles aleatorizados, sintéticos y de series temporales.

Referencias

Chevallard, Yves (1994). "Les processus de transposition didactique et leur théorisation". ARSAC

Clow, D. (2012). The learning analytics cycle: closing the loop effectively.

G. ET ALII (COORD.) (1994). "La transposition didactique à l'épreuve" Paris, La Pensée, Sauvage, 135 – 1

Hanusheck, Eric (1981). "Throwing Money at Schools". Journal of Policy Analysis and Management.

Stanley, Donald y Campbell, Julian (1963). "Experimental and quasi-experimental designs for research". Houghton Mifflin Company Boston.