

Disorder Index for Identifying Students with Academic Delay

Jewsbury Alejandra, García Mario Alejandro, Giubergia Fernanda, Socolovsky Silvia, and Guzmán Analía

Universidad Tecnológica Nacional, Facultad Regional Córdoba
ajewsbury@gmail.com

Abstract This work presents preliminary results of analyzing academic data to identify student populations showing delays or academic setbacks in engineering programs, based on previous hypotheses from the team's research regarding student cohort behavior and academic progress. Alongside research using data from university information systems, qualitative studies contribute a key aspect for review: student behavior throughout their academic journey influences degree retention. This presentation introduces the “disorder index”, a term coined by the team that reflects what is described in the literature as timely progression. Through this index, large volumes of data are analyzed, strengthening the hypothesis that academic delay conditions university persistence.

Keywords: dropout , education delay

Indicador de Desorden en la Identificación de Población Estudiantil en Situación de Retraso Académico

Resumen Este trabajo presenta los resultados preliminares del análisis de datos académicos para identificar población estudiantil que presenta atrasos o rezagos en el desarrollo académico en carreras de ingeniería atendiendo hipótesis previas resultantes de investigaciones del equipo en referencia al comportamiento de las cohortes de estudiantes y su desarrollo académico. A la investigación con datos provenientes de sistemas informáticos de la universidad se le asocian estudios cualitativos que aportan un aspecto a revisar: la permanencia en la carrera no es independiente del comportamiento del estudiante en el trayecto académico. En esta presentación se introduce el indicador de desorden, término acuñado en el equipo, que refleja lo que en la literatura se denomina cursado en tiempo y forma. Entonces, por medio del desorden es que se analizan los

datos disponibles y se permite fortalecer la hipótesis de que el retraso académico condiciona la permanencia en la universidad.

Keywords: abandono, retraso académico

1. Contexto

El trabajo se desarrolla en el marco del proyecto PID UTN 533, titulado “Predicción del riesgo de abandono de estudiantes en las carreras de Ingeniería de la Facultad Regional Córdoba de la UTN”. Este proyecto, iniciado en el año 2025, cuenta con la participación de investigadores de los grupos GIEIMET (Enseñanza de la Ingeniería Mediada por Tecnología) y GIA (Grupo de Investigación en Inteligencia Artificial) de la misma facultad. El objetivo principal de este proyecto de investigación es obtener una valoración de riesgo de abandono individualizada para cada estudiante mediante modelos de aprendizaje automático factibles de construir a partir de las fuentes de datos institucionales. No obstante, este trabajo solo avanza a modo preliminar en la discusión resultante de análisis estadísticos de datos del sistema académico que reflejan la tendencia de que el retraso en los estudios afecta a la graduación a término y que podría evidenciar, con estudios que aún no contamos, como indicador de abandono. La referencia conceptual en la literatura, discusión sobre el concepto de abandono es extensa; a menudo se distingue entre el abandono de una carrera y el de una asignatura específica, y se diferencia del término "deserción" según las acciones de retención asumidas. En los aspectos de la economía de la educación rescatamos que los estudios sistemáticos cobraron relevancia en América Latina a partir del año 2000, vinculados a contextos de cuestionamiento de fondos de financiamiento internacionales en referencia al gasto público del estado en el sostenimiento del sistema de educación superior. Ante la diversidad de definiciones de abandono este equipo adquiere la definición de no continuidad que surge de la Real Academia Española y que adopta el Ministerio de Educación de la Nación: un estudiante abandona cuando al año siguiente no se matricula (Ministerio de Capital Humano, 2025).

2. Introducción

Este trabajo presenta un análisis preliminar sobre los conceptos de abandono y retraso para las carreras de ingeniería de la UTN-FRC, integrando el marco conceptual de la semántica del desarrollo académico curricular en bases a ciclos, propios de los trayectos formales de educación, con la elaboración de modelos estadísticos predictivos sobre una colección completa de datos académicos de la población estudiantil. El análisis de la literatura sobre la temática de abandono en universidades ha permitido elaborar un conjunto de definiciones que dan marco al trabajo actual y a futuras investigaciones del equipo en el proyecto antes citado. Las variables definidas son: duración de la carrera, retraso, terminalidad a término y abandono. En este momento el presente trabajo solo analiza retraso

y desarrollo académico y lo referencia anticipatoriamente con abandono, tratándose entonces de hallazgos preliminares para la investigación total del proyecto (Jewsbury et al., 2025).

A referencia de marco de relevancia de la temática que se aborda se indica que la situación de abandono en las universidades se presenta como sistémica en todo el mundo y no escapa a instituciones privadas o públicas. Una oportunidad presente hoy es que la tecnología dispone de medios innovadores de exploración y análisis de datos lo que facilita la comprensión de este fenómeno social. Este proyecto de investigación en curso persigue el objetivo de anticipar situaciones de riesgo de abandono para activar acciones de reparación efectivas. Identificamos entonces que apelamos a acciones de alerta temprana, en el proyecto de investigación que da marco a este trabajo, en cuanto se identifica oportunamente a estudiantes vulnerables para intervenir antes de que se produzca el abandono (Donoso-Díaz et al., 2018).

3. Objetivos

Los objetivos del trabajo son:

- Presentar y analizar los resultados de la aplicación de modelos estadísticos sobre datos académicos observando situaciones de retraso en función de la aprobación en desorden de materias de distintos ciclos académicos de una carrera universitaria.
- Organizar la información recolectada en el marco semántico de los estudios de abandono.
- Validar la hipótesis de que el desorden en la forma de rendir las materias es una variable asociada al retraso y al abandono.

4. Desarrollo

El estudio del marco teórico de la problemática ha permitido seleccionar un conjunto de características de representación de abandono en las universidades. Entre estas se encuentra el concepto de terminalidad a tiempo, resultante de adicionar un año a los años de cursado, esto expresa para las carreras de ingeniería una terminalidad a término de 6 años. En este marco, y analizando los datos del sistema académico en otra investigación de este equipo para las cohortes 2017 hasta 2024, los estudiantes que completan sus estudios a término es 23 % del total de inscriptos en la misma cohorte.

Entonces quienes no se encuentran comprendidos en la terminalidad a término tendrán retraso en su egreso o han abandonado. Y por tanto observando los exámenes finales aprobados, para toda la población, se propone el estudio del retraso académico.

4.1. El retraso

El retraso en los estudios se refiere a aquellas trayectorias académicas que no siguen el ritmo esperado según el plan de estudios, superando la terminalidad a término. Este fenómeno se traduce en un desfase temporal en el avance del estudiante, asociada a la acumulación de materias no aprobadas y a una mayor permanencia en el sistema, constituyéndose en una variable relevante en el análisis del abandono (Moreno et al., 2019).

Investigaciones previas en base al relevamiento a estudiantes que han abandonado los estudios explican que el abandono tiene en mayor porcentaje de ocurrencia causas radicadas en dificultades administrativas y académicas. Indicadas como causas administrativas (modalidad de cursado, régimen de examen final, documentación y accesibilidad en la información) y como causas académicas (no regularidad de las materias, no aprobación de exámenes finales, no comprensión de contenidos y prácticas), agregando una categoría de causas académico-administrativas asociadas a la organización del currículo (cantidad, tipo y carga horaria de materias por ciclo, régimen de correlatividad, titulación intermedia).

El cursado en tiempo y forma

Se llama cursado en tiempo y forma al trayecto que sigue un estudiante durante un estudio formal siguiendo los lineamientos del plan de estudio correspondiente. Este concepto es utilizado en la educación secundaria en distintas investigaciones observando aquellos estudiantes que no repiten ciclo, que no rinden exámenes, ni exámenes previos. Es decir, estudiantes que completan los estudios en el tiempo establecido y sin utilizar las instancias de exámenes regulares ni previos (Kit et al., 2023). Este concepto no puede trasladarse directamente al sistema universitario ya que no obstante encontrar ciclos los estudiantes al finalizar el cursado, ya sea en condición de libre, promoción o regular rinden exámenes finales. También, los planes de carrera, presentan las asignaturas por ciclo pero cada asignatura puede o no tener un conjunto de asignaturas de requerimiento previo. Esto lleva a que un estudiante puede realizar el recorrido permitido en el plan movilizándose entre los ciclos.

4.2. Construcción del concepto desorden

Considerando el concepto de cursado en tiempo y forma se introduce una nueva variable de análisis al comportamiento de estudiantes en estudios universitarios, este es el concepto de desorden o no cumplimiento del plan ciclo por ciclo. La variable de desorden de la trayectoria académica se define como una medida que cuantifica el grado en que la aprobación de asignaturas se aparta o es distinto al orden establecido en el plan de estudios. Para esta investigación se considera esta característica como un indicador que toma valores entre 0 y 1. Donde 0 indica estudiantes que desarrollan su carrera aprobando las materias de nivel a nivel siguiente, es decir, completando un nivel antes de pasar al

siguiente. El índice de desorden de trayectoria académica permite capturar desviaciones en la secuencia esperada de avance curricular, constituyéndose como una medida complementaria al retraso en los estudios. Mientras que el retraso mide el tiempo que le lleva al estudiante completar la carrera, el desorden describe la forma en que ese avance ocurre, poniendo en evidencia situaciones de eventual retraso o quizás abandono. Cabe señalar que el indicador de desorden no expresa la forma en la que el estudiante aprueba las materias con relación al régimen de correlativas, solo lo indica en función de las asignaturas que aprueba correspondientes a un nivel.

Método de cálculo del grado de desorden Para el cálculo de la variable desorden, se considera un vector ordenado cronológicamente que contiene cada uno de los niveles de las asignaturas aprobadas por el estudiante. Se define la variable u que va almacenando el último nivel alcanzado, y un contador de desorden c inicializado en cero. Inicialmente, u toma el valor del nivel de la primera asignatura aprobada. Luego, se recorre el vector comparando cada nivel con el valor almacenado en u :

- Si el nivel leído es mayor o igual a u , no se registra desorden y, en caso de ser mayor, se actualiza el valor de u .
- Si el nivel leído es menor que u , se registra un evento de desorden incrementando el contador c en una unidad.

Finalmente, la variable de desorden se calcula como el cociente entre el contador c y la cantidad total de asignaturas aprobadas por el estudiante.

5. Metodología

Los datos utilizados en la elaboración de este primer modelo predictivo corresponden a datos del historial académico de estudiantes de la Facultad Regional Córdoba provenientes del sistema académico institucional (SisAcad). Los datos colocados a disposición de este proyecto de investigación son: Exámenes (carrera, materia, fecha y nota), materias y planes de estudio.

La metodología que se ha seguido incluye las siguientes etapas:

- Extracción y anonimización. Los datos originales (SQL Server) se han incorporado a una base de datos SQLite para facilitar su manejo, eliminando datos sensibles como CUIL, correos y teléfono.
- Selección de la muestra. Selección inicial de un subconjunto en base a una totalidad de estudiantes ingresantes entre los años 1985 y 2017. No se consideran personas con ingreso a partir del año 2018 para acotar el análisis a tres planes de estudio específicos y para que sea válido el criterio, que se explica más adelante, utilizado para definir la situación de abandono.
- Ingeniería de características. Introducción de la variable desorden.
- Validación estadística. Planteo de la hipótesis: desorden medio es significativamente mayor en los casos de abandono.
- Análisis de resultados. Valoración de las respuestas obtenidas en función del marco teórico construido alrededor de la temática de abandono.

5.1. Variable *abandono*

El abandono, asociado por definición a quienes no se matriculan en el ciclo siguiente, no se analizó dado que no se cuenta con información (en este conjunto de datos) sobre la matrícula año a año del estudiante. Por esta razón se ha construido con una aproximación identificando la cantidad de materias aprobadas. Para la variable *abandono* se considera entonces, que un estudiante que tiene más de 38 materias aprobadas ha egresado. En caso contrario, y dado que el año mínimo de ingreso es 2017, se considera abandono. En la Figura 1 se muestra la distribución de la cantidad de materias aprobadas para el universo de estudiantes considerados. Notar que existe un máximo local para el valor 43, y que la tendencia alista comienza en 39. Es lógico pensar que esta moda secundaria se debe a la acumulación de materias aprobadas en las personas que concluyen satisfactoriamente sus estudios.

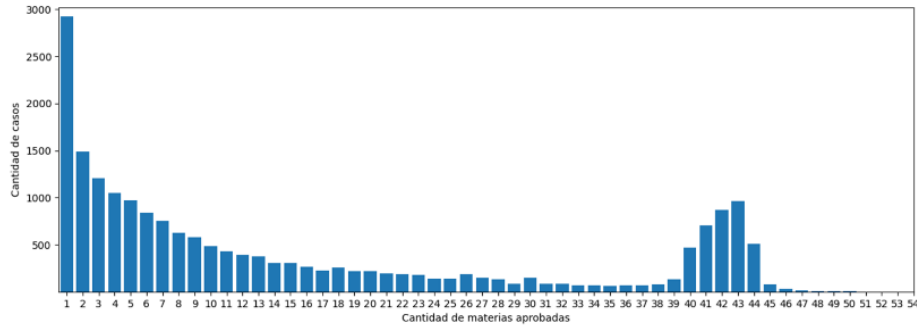


Figura 1. Distribución de la cantidad de materias aprobadas (Período 1985-2017)

5.2. Variable *desorden*

La probabilidad de rendir materias fuera de orden aumenta con la cantidad de materias aprobadas, por lo tanto, las trayectorias más largas tienden a tener mayor grado de desorden. Para que sea posible comparar trayectorias de distintas longitudes, se definen 7 versiones del indicador según la cantidad de materias consideradas:

- Desorden_05: desorden de las primeras 5 materias aprobadas
- Desorden_10: desorden de las primeras 10 materias aprobadas
- Desorden_15: desorden de las primeras 15 materias aprobadas
- Desorden_20: desorden de las primeras 20 materias aprobadas
- Desorden_25: desorden de las primeras 25 materias aprobadas
- Desorden_30: desorden de las primeras 30 materias aprobadas
- Desorden_35: desorden de las primeras 35 materias aprobadas

Algunas consideraciones sobre la variable *desorden*:

- No se puede calcular el desorden sobre estudiantes que aprobaron menos de 2 materias.
- El desorden será cero siempre que todas las materias aprobadas pertenezcan a primer año.

5.3. Prueba de hipótesis

Para probar la hipótesis de que el abandono no es independiente del desorden, se plantea la hipótesis de que las medias del desorden son iguales para los casos de abandono y no abandono. La refutación de la segunda hipótesis prueba la segunda. La comparación de las medias se llevó a cabo por ambientes. Se consideraron 21 ambientes, conformados por 21 conjuntos de datos diferentes. Los ambientes se definen según las combinaciones de los siguientes criterios:

- Plan de estudio (plan 85, plan 95, plan 2008)
- Cantidad de materias aprobadas (>4, >9, >14, >19, >24, >29, >34) Para cada ambiente se realizó la comparación de medias de las 7 variantes de la variable desorden, resultando en 147 pruebas.

6. Resultados

Los resultados son consistentes para todos los ambientes (salvo pocas excepciones que se comentan más adelante). El desorden medio difiere entre los estudiantes que terminan sus estudios y quienes abandonan. Los estudiantes con mayor desorden tienen probabilidades más altas de abandonar. Las excepciones ocurren solo en ambientes extremos, por ejemplo cuando se utiliza el desorden_35 con el grupo de estudiantes que aprobó más de 4 materias. En esos casos, el gran volumen de estudiantes que aprobaron pocas materias y luego abandonaron decrementa el desorden medio porque, como se mencionó antes, el desorden crece con la cantidad de materias aprobadas.

A modo de ejemplo, en la Figura 2 se muestra una comparación de las distribuciones del desorden_25 para estudiantes del plan 85 con 10 materias aprobadas o más. A continuación, para el mismo ambiente, se pueden ver los datos calculados:

- Desorden medio para abandono: 0.2212
- Desorden medio para no-abandono: 0.1313
- $t = 16.912$
- p-valor: 1.0956
- Diferencia de medias: 0.0899
- IC 95 %: 0.079, 0.1

Para este ambiente, el test arroja un p-valor que descarta que la diferencia sea producto de la casualidad y confirma al desorden como una característica relevante para la predicción. Tal como se comentó, salvo para ambientes extremos los resultados son similares en el resto de los ambientes.

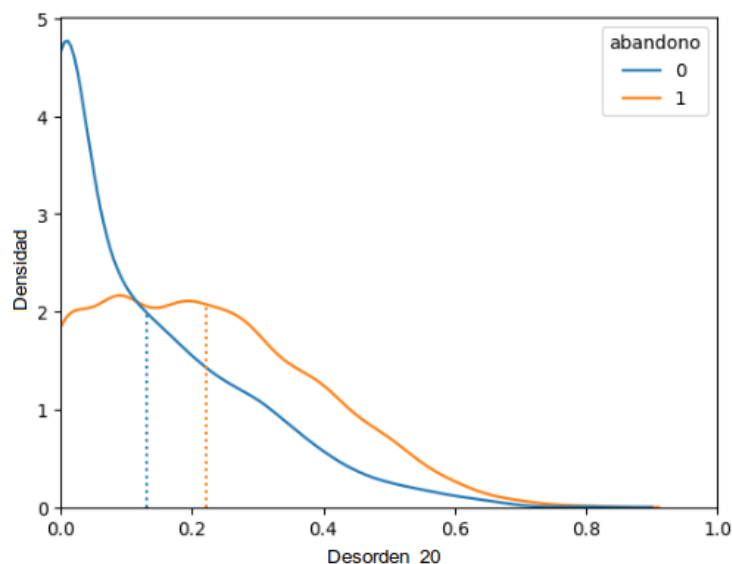


Figura 2. Densidad de probabilidad (línea continua) de `desorden_20` según el abandono para estudiantes con 10 o más materias aprobadas en plan 85. En línea de puntos se muestran las medias de cada distribución.

7. Conclusiones y próximos pasos

La investigación confirma que es posible identificar patrones académicos que reflejan desorden en la forma en la que se desarrollan los exámenes, que esta definición se asocia al retraso y que es posible, en futuras investigaciones, reflejar como abandono mediante análisis estadístico. La variable de desorden en la trayectoria académica ha demostrado ser un predictor potente y válido. Esta primera inspección no solo expresa el comportamiento de la población estudiantil en el recorrido académico sino también brinda información para analizar los diseños curriculares, permitiendo ponderar la dificultad de las asignaturas de un mismo nivel, o agrupamiento de asignaturas de bloques poco motivadores en un mismo nivel, o de bloques de materias de mayor exigencia de estudios que funcionen como “cuello de botella”. El diseño curricular se evidencia entonces como el espacio donde transcurre la organización del estudiante donde el cursado en tiempo y forma (de ciclos) se correlaciona con la garantía de egreso. Este trabajo inicia, para el grupo de investigación, el interrogante de indagar como el desorden se relaciona con los bloques temáticos de la disciplina o con determinados niveles o asignaturas, permitiendo revisar el diseño curricular de la carrera o administrar recursos en acciones reparadoras.

Declaración de conflictos de intereses Los autores declaran no tener conflictos de intereses que sean relevantes para el contenido de este artículo.

Referencias

- Donoso-Díaz, S., Iturrieta, T. N., & Traverso, G. D. (2018). Sistemas de Alerta Temprana para estudiantes en riesgo de abandono de la Educación Superior. *Ensaio: Avaliação e Políticas Públicas em Educação*, 26(100), 944-967.
- Jewsbury, A., Giubergia, F., & Socolovsky, S. (2025). Aproximación al estudio del abandono. *Congreso Nacional de Sistemas de Información*.
- Kit, I., Nistal, M., & Sáenz Guillén, L. (2023). Índice de Resultados Escolares: ¿Cuántos estudiantes llegan al final de la secundaria en tiempo y forma. *Observatorio de Argentinos por la Educación*.
- Ministerio de Capital Humano. (2025). *Anuario Estadístico de Educación Superior 2023-2024: Síntesis de información*. Ministerio de Capital Humano.
- Moreno, J. E., Chiecher, A., & Paoloni, P. (2019). Los estudiantes universitarios y sus metas académicas: Implicancias en el logro y retraso de los estudios. *Ciencia, docencia y tecnología*, (59), 148-173.