

ANÁLISIS DE ALGUNAS VARIABLES PARA DETERMINAR PERFILES DE DESEMPEÑO ACADÉMICO EN LA UNNE

Julio César Acosta

Facultad de Ciencias Exactas y Naturales
y Agrimensura - Facultad de Ciencias
Agrarias. Universidad Nacional del Nordeste
Corrientes, Argentina.

David Luis La Red Martínez

Facultad de Ciencias Exactas y
Naturales y Agrimensura. Universidad
Nacional del Nordeste
Corrientes, Argentina.

All content in this magazine is
licensed under a Creative Com-
mons Attribution License. Attri-
bution-Non-Commercial-Non-
Derivatives 4.0 International (CC
BY-NC-ND 4.0).



Resumen: En este trabajo se propone evaluar variables que pueden incidir en el rendimiento de los estudiantes. La propuesta no se enfoca en analizar el perfil del estudiante solo a través de sus calificaciones, sino también, estudiar el desempeño académico en base a otras variables que puedan determinar patrones que conduzcan al éxito o fracaso académico, en este trabajo presentamos los resultados de las variables: sexo; lugar de procedencia del alumno y tiempo dedicado al estudio. Trabajamos en implementar un modelo que relacione las calificaciones de los estudiantes con las variables mencionadas y otras, tales como factores socioeconómicos, demográficos, actitudinales, entre otros; lo cual servirá para clasificar los diferentes perfiles de alumnos. Los datos relevados corresponden a alumnos de la asignatura Matemática I de la carrera Ingeniería Agronómica (IA) de la Facultad de Ciencias Agrarias (FCA) de la Universidad Nacional del Nordeste (UNNE) cohorte 2017 en el marco del PI 16F002 acreditado por Res. N° 970/16 CS y fueron procesados con el uso de InfoSphere Warehouse. Esperamos contribuir a encontrar una respuesta al bajo rendimiento académico de los alumnos observado históricamente, problema éste que es el disparador de nuestra investigación. Los modelos predictivos que buscamos, permitirán tomar acciones tendientes a evitar el fracaso académico, detectando los alumnos con perfil de riesgo de fracaso académico de manera temprana, a poco del inicio del cursado de las asignaturas, lo que permitirá concentrar en ellos los esfuerzos de tutorías y apoyos especiales.

Palabras clave: Rendimiento académico, almacenes de datos, minería de datos, modelos predictivos.

INTRODUCCIÓN

En nuestra Universidad, que tiene políticas definidas de inclusión y contención, uno de

los factores que evaluamos continuamente es el rendimiento académico. Se entiende el rendimiento académico como la productividad del sujeto, matizado por sus actividades, rasgos y la percepción más o menos correcta de los cometidos asignados (MALETIC, COLLARD y MARCUS, 2002). Presentamos valores de las variables a) Sexo; b) Provincia de procedencia y c) Tiempo dedicado al estudio, las que serán analizadas para determinar si pueden ser o no identificadas como incidentes en el rendimiento académico.

La Minería de Datos (del inglés Data Mining - DM), son procesos de descubrimiento de nuevas y significativas relaciones, patrones y tendencias en grandes volúmenes de datos utilizando técnicas de Artificial intelligence (AI) y Machine learning (ML). Estas técnicas permiten extraer patrones y tendencias para describir y comprender mejor los datos y predecir comportamientos futuros. Un DW es una colección de datos orientada a un dominio, integrada, no volátil y variante en el tiempo para ayudar a tomar decisiones (SALTON, 1989). Los DW surgieron por dos razones: a) la necesidad de proporcionar una fuente única de datos limpia y consistente para propósitos de apoyo para la toma de decisiones; b) la necesidad de hacerlo sin afectar a los sistemas operacionales (MOLINA LOPEZ y GARCÍA HERRERO, 2006). En este trabajo, que se encuentra en desarrollo, se utilizaron de técnicas de DM, con volúmenes no muy grandes de datos, sobre la información del desempeño de los alumnos de la cátedra Matemática I de la FCA-UNNE. Describimos el contexto de la experiencia y algunos resultados preliminares.

MATERIALES Y MÉTODOS

Los resultados de la experiencia que exponemos en este trabajo pertenecen a la asignatura Matemática I de la carrera IA de la FCA de la UNNE que se dicta en el trimestre

marzo, abril y mayo y tiene una carga horaria de 96 (noventa y seis) horas reloj de dictado presencial con una distribución de tiempos de dictado de teoría y de trabajos prácticos del 50% para cada una.

Para alcanzar la condición de alumno regular, los alumnos deben asistir al menos al 75% de las clases de trabajos prácticos, que se dictan dos veces por semana en clases de 2 hs. cada una y deben aprobar 2 (dos) exámenes parciales cuyos contenidos son exclusivamente de trabajos prácticos; cada uno de ellos tiene su instancia de recuperación y para aquellos alumnos que hayan aprobado al menos 1 (uno) de los parciales en cualquiera de las 4 (cuatro) instancias disponibles, existe una instancia más para recuperar el examen que queda aún sin aprobar. Cualquiera de los exámenes parciales se aprueba con 60 (sesenta) puntos sobre 100 (cien) puntos posibles. La asistencia a clases de teoría es libre y se dictan dos veces por semana en clases de 2 hs. cada una.

Se acredita la asignatura con un examen final al que se accede en condición de alumno regular o de alumno libre; el alumno regular debe rendir en el examen final solamente los contenidos de teoría en un examen oral. El alumno que se presenta al examen final en condición de alumno libre, debe rendir un examen escrito de trabajos prácticos y tras aprobar esa instancia pasa al examen de teoría en condiciones similares a la antes mencionada.

La cantidad de alumnos que regularizan y/o que aprueban las asignaturas involucradas en este proyecto no es satisfactoria, consideramos que esa situación puede contribuir al desgranamiento y deserción de los alumnos en los primeros niveles de sus carreras. Es importante, por tanto, estudiar y determinar cuáles son las variables que inciden en el rendimiento académico a fin de poder establecer estrategias de acción pedagógicas que permitan mejorar dicho rendimiento.

Trabajamos principalmente en el desarrollo de métodos que contribuyan a encontrar técnicas para la detección temprana de los alumnos que tendrán dificultades en sus estudios, a fin ofrecerles una contención y acompañamiento especial en el inicio de sus estudios Universitarios.

Con la finalidad de adoptar acciones proactivas frente al desgranamiento y el bajo rendimiento académico de los alumnos nos proponemos detectar grupos de estudiantes en riesgo de fracaso en sus estudios, a partir de determinadas variables, de las cuales en este trabajo nos centramos en tres de ellas: a) Sexo, b) Provincia de procedencia, c) Horas dedicadas al estudio. Hemos asignado a cada una de las variables las siguientes sub-variables: para Sexo: Masculino – Femenino; para Provincia de procedencia: Corrientes, Chaco y Otras provincias ya que la FCA también recibe alumnos que proceden de las Provincias de Formosa, Misiones, Santiago del Estero, Salta y Jujuy, todas ellas ubicadas en un radio de 800 km. de la sede de la FCA en la ciudad de Corrientes, Capital de la Provincia del mismo nombre y Horas dedicadas al estudio: Hasta 10 hs, de 10 hs. a 20 hs., más de 20 hs. y no sabe / no contesta. Las variables y subvariables en cuestión fueron analizadas discriminando el resultado final del cursado: en alumnos Regulares (R), alumnos Libres por parciales (LPP), alumnos Libres por inasistencias (LPI) y alumnos matriculados que nunca asistieron (N/A). En las figuras que siguen se presentan los valores resultantes de la cohorte 2019.

Hemos trabajado con técnicas de Data Warehouse (DW), que es un sistema especial de bases de datos utilizado para el almacenamiento de datos y el procesamiento de los mismos para la presentación de informes y análisis de información, es considerado como un componente central de la inteligencia de organizaciones. Un DW es un repositorio

de datos que proporciona una visión global, común e integrada de los datos (CURTO DIAZ, 2010). Existen diversos estudios y publicaciones que abordan la evaluación de rendimiento académico utilizando técnicas de Minería de Datos (FORMIA y LANZARINI, 2013); (PEREIRA, ROMERO y TOLEDO, 2013); (LA RED MARTINEZ, GIOVANNINI, BAEZ MOLINAS, TORRE y YACCUZZI, 2017).

Datos obtenidos considerando la variable sexo

En Figuras 1 a 8 se presentan los valores obtenidos para la variable sexo, analizados desde diferentes puntos de vista.

Empezamos presentando en la Figura 1 los porcentuales totales de varones y mujeres inscriptos en la corte 2018 sobre un total de 206 alumnos. En la Figura 2 se presentan los porcentuales de la discriminación de la totalidad de los alumnos en Regulares, Libres por parciales (LPP), Libres por inasistencia (LPI) y Nunca asistieron (N/A).

En la Figura 3 se presenta la misma información que en la Figura 2 referida a los alumnos de sexo masculino, por tanto la sumatoria de los porcentuales del gráfico totalizan el 59,59% que es el porcentual total de alumnos de sexo masculino y en la Figura 4 ídem referido a las alumnas de sexo femenino por tanto la sumatoria de los porcentuales volcados al gráfico totaliza el 40,41%.

Cuando comparamos las Figuras 3 y 4 observamos que el porcentual total de alumnos Regulares del sexo masculino del 32,31%, indica que un 54,22% del total de alumnos del sexo masculino regularizaron la asignatura; mientras en el caso de las alumnas del sexo femenino regularizaron la asignatura el porcentual total del 26,43%, indica que un porcentual relativo del 65,41% del total de las alumnas regularizaron la asignatura; lo que arroja una diferencia de porcentuales relativos

de 11,19% a favor del total relativo femenino.

Esta diferencia no es apreciable en el caso de los alumnos libres por parciales ya que el porcentual total de 12,23%, indica que un 20,52% del total de alumnos quedaron libres por parciales; mientras en el caso de las alumnas el porcentual del 8,16%, indica que un porcentual relativo del 20,19% del total de las alumnas quedaron libres por parciales en la asignatura, lo que da una pequeña diferencia de porcentuales relativos totales del 0,33% a favor del total relativo masculino.

En el caso de los alumnos Libres por inasistencias, el porcentual total de 8,22%, indica un 13,80% del total relativo de alumnos del sexo masculino que quedaron libres por inasistencias en la asignatura; mientras que en el caso de las alumnas el porcentual total del 2,46%, indica un porcentual total relativo del 6,09% del total de las alumnas que quedaron libres por inasistencias en la asignatura, con una diferencia del orden de 7,71% a favor del total relativo masculino.

Para la categoría alumnos que nunca asistieron el porcentual total de 6,83%, indica un 11,46% del total de alumnos del sexo masculino que quedaron libres por inasistencias en la asignatura (total relativo masculino); mientras que en el caso de las alumnas del sexo femenino el porcentual total del 3,36%, indica un porcentual relativo del 8,31% del total de las alumnas de sexo femenino que quedaron libres por inasistencias en la asignatura (total relativo femenino), con una diferencia del orden de 3,15%.

Sometemos estos valores a una prueba del test del Chi cuadrado que arroja los siguientes valores y conclusiones:

El Chi de tabla está obtenido para un nivel de confiabilidad del 95%, encontramos que al ser el Chi calculado = 4,19029139 < Chi de tabla = 7,8147279; las diferencias de valores obtenidos entre resultados observados y los esperados (de regulares y no regulares) son

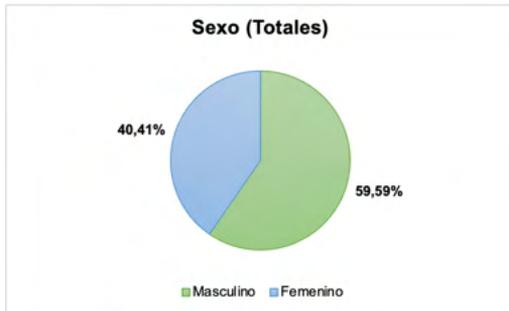


Figura 1: Porcentuales totales de alumnos por sexo.



Figura 2: Porcentuales totales de alumnos por los resultados obtenidos.

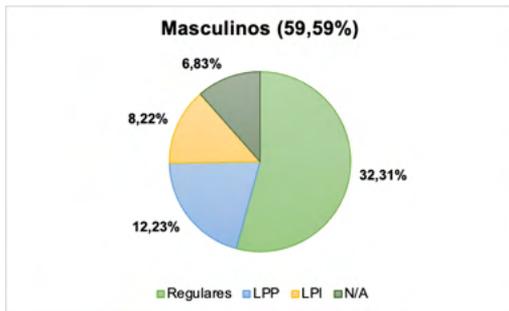


Figura 3: Porcentuales de varones por los resultados obtenidos.

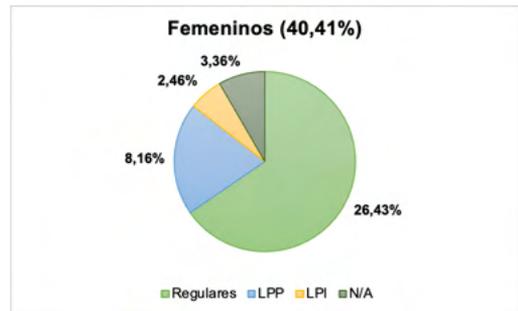


Figura 4: Porcentuales de mujeres por los resultados obtenidos.

	Regulares	LPP	LPI	N/A	Totales
M	67	25	17	14	123
F	54	17	5	7	83
Totales	121	42	22	21	206

Tabla 1: Valores observados discriminados por sexo

	Regulares	LPP	LPI	N/A
M	72,2475728	25,0776699	13,1359223	12,538835
F	48,7524272	16,9223301	8,86407767	8,46116505
Totales	121	42	22	21

Tabla 2: Valores esperados discriminados por sexo

Chi calculado	4,19029139
Chi de tabla	7,8147279

Tabla 3: Valores de Chi

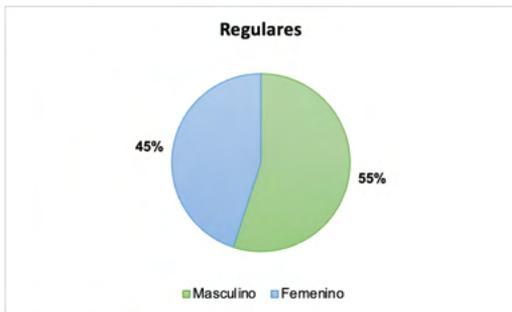


Figura 5: Porcentuales de alumnos Regulares por sexo.



Figura 6: Porcentuales de alumnos Libres por parciales por sexo.



Figura 7: Porcentuales de alumnos Libres por inasistencias por sexo.



Figura 8: Porcentuales de alumnos que nunca asistieron por sexo.

aceptables dentro del margen de lo aleatorio y no se deben a la influencia de la variable analizada, la que es en este caso el sexo.

Las Figuras 5, 6, 7 y 8 presentan respectivamente los porcentuales de alumnos Regulares, Libres por parciales, Libres por inasistencias y que nunca asistieron.

Datos obtenidos considerando la variable procedencia

En las Figuras 9 a 16 se presentan los valores obtenidos para la variable procedencia, analizados desde diferentes puntos de vista.

En la Figura 9 presentamos los porcentuales totales discriminados por provincias de procedencia de los 206 alumnos inscriptos en la cohorte 2018. En la Figura 10 se presentan los porcentuales de la totalidad de los alumnos de la provincia de Corrientes discriminados en Regulares, Libres por parciales, Libres por inasistencia y que nunca asistieron.

En la Figura 11 se presenta la misma información que en la Figura 10 pero referida a los alumnos procedentes de la Provincia de Chaco y en la Figura 12 referida a los alumnos procedentes de Otras Provincias.

Cuando analizamos las Figuras 10, 11 y 12 observamos que el porcentual total de alumnos regulares procedentes de la Provincia de Corrientes del 34,07%, indica que un 54,80% del total de alumnos procedentes de la Provincia de Corrientes regularizaron la asignatura, mientras que en el caso de los alumnos procedentes de la Provincia de Chaco que regularizaron la asignatura el porcentual total del 19,38%, lo que indica que un porcentual relativo del 64,41% del total de los alumnos de la Provincia de Chaco regularizaron la asignatura y de los valores de los alumnos de Otras Provincias que regularizaron la asignatura se tiene que el porcentual total del 5,29% indica que un porcentual relativo del 68,35% del total de

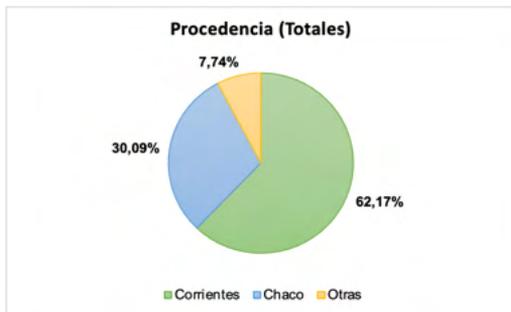


Figura 9: Porcentuales totales de alumnos por procedencia.

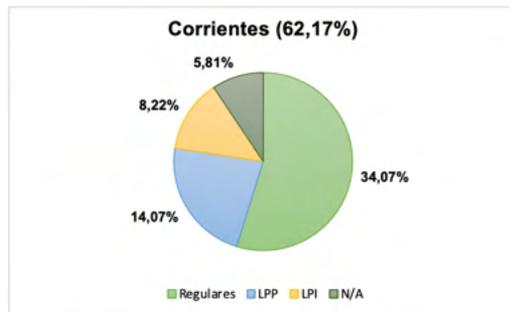


Figura 10: Porcentuales totales de alumnos de Corrientes por resultados.

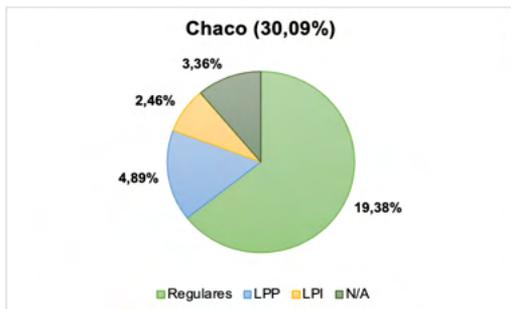


Figura 11: Porcentuales totales de alumnos de Chaco por resultados.

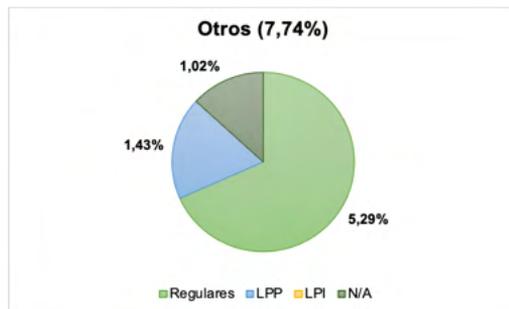


Figura 12: Porcentuales totales de alumnos de Otros lugares por resultados.

	Regulares	LPP	LPI	N/A	Totales
Corrientes	70	29	17	12	128
Chaco	40	10	5	7	62
Otros	11	3		2	16
Totales	121	42	22	21	206

Tabla 1: Valores observados discriminados por sexo

	Regulares	LPP	LPI	N/A
Corrientes	75,184466	26,0970874	13,6699029	13,0485437
Chaco	36,4174757	12,6407767	6,62135922	6,32038835
Otros	9,39805825	3,26213592	1,70873786	1,63106796
Totales	121	42	22	21

Tabla 2: Valores esperados discriminados por provincia de procedencia

Chi calculado	5,03641965
Chi tabla	12,5915872

Tabla 3: Valores de Chi

los alumnos procedentes de Otras Provincias regularizaron la asignatura, resultando de esto una diferencia de porcentuales relativos de 13,55% entre los alumnos procedentes de Corrientes y de Otras provincias.

Esta diferencia se reduce en el caso de los alumnos libres por parciales ya que el porcentual total de alumnos procedentes de Corrientes del 14,07%, indica un 22,63% del total de alumnos procedentes de esa provincia que quedaron libres por parciales en la asignatura; en el caso de alumnos procedentes de la provincia de Chaco el 4,89%, indica que un 16,25% del total de alumnos procedentes de esa provincia quedaron libres por parciales en la asignatura y para los alumnos procedentes de Otras provincias se registró un 1,43% en total para esta subvariable, que resulta un 18,48% del total de alumnos procedentes de Otras provincias que quedaron libres por parciales en la asignatura; resultando así una diferencia entre el mayor valor y el menor valor de 6,38%.

En alumnos libres por inasistencias el porcentual total de alumnos de Corrientes del 8,22%, indica que un 13,22% del total de alumnos procedentes de esa provincia quedaron libres por inasistencias en la asignatura; para los alumnos de la provincia de Chaco el porcentual total fue 2,46%, que indica que un 8,18% del total de alumnos procedentes de esa provincia quedaron libres por inasistencias en la asignatura y para los alumnos de Otras provincias no se registraron alumnos libres por inasistencia en la asignatura; resultando así una diferencia entre el mayor valor y el menor valor del 13,22%.

Para los alumnos que nunca asistieron el porcentual total de alumnos procedentes de Corrientes fue 5,81%, que indica un 9,35% del total de alumnos procedentes de esa provincia que nunca asistieron; en el caso de alumnos procedentes de la provincia de Chaco el 3,36%, indica un 11,17% del total de alumnos

procedentes de esa provincia que nunca asistieron y para los alumnos procedentes de Otras provincias se registró un 1,02% en total, que resulta un 13,18% del total de alumnos de Otras provincias que nunca asistieron; resultando así una diferencia entre el mayor valor y el menor valor de 3,83%.

Sometemos entonces estos valores a una prueba del test del chi cuadrado que arroja los siguientes valores y conclusiones:

El Chi de tabla está obtenido para un nivel de confiabilidad del 95%, encontramos que al ser el Chi calculado = 5,03641965 < Chi de tabla = 12,5915872, indica que las diferencias de valores obtenidos entre resultados observados y los esperados son aceptables dentro del margen de lo aleatorio y no se deben a la influencia de la variable analizada, la que es en este caso la provincia de procedencia.

En las Figuras 13, 14, 15 y 16 se presentan respectivamente los porcentuales de alumnos Regulares, Libres por parciales, Libres por inasistencias y que nunca asistieron discriminados por provincias de procedencia.

Datos obtenidos considerando la variable horas dedicadas al estudio

En las Figuras 17 a 20 se presentan los valores obtenidos para el tiempo dedicado al estudio, analizados desde diferentes puntos de vista.

En la Figura 17 presentamos los porcentuales totales discriminados por tiempo destinado al estudio de los 206 alumnos inscriptos en la cohorte 2018. En la Figura 18 se presentan los porcentuales de la totalidad de los alumnos que dedican hasta 10 hs. semanales a sus estudios discriminados en Regulares, Libres por parciales, Libres por inasistencia y que nunca asistieron. En la Figura 19 se presenta la misma información que en la Figura 18 pero referida a los alumnos que dedican entre 10 hs. y 20 hs. semanales a sus estudios y en la

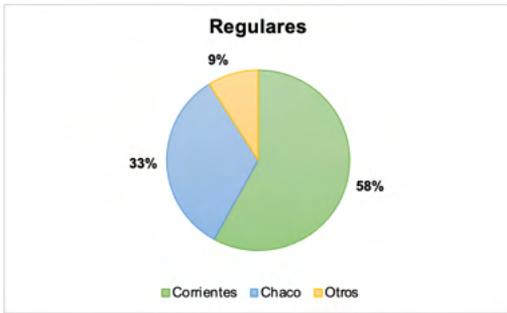


Figura 13: Porcentuales de alumnos regulares por procedencia.



Figura 14: Porcentuales de alumnos Libres por parciales por procedencia.



Figura 15: Porcentuales de alumnos Libres por inasistencias por procedencia.



Figura 16: Porcentuales de alumnos que nunca asistieron por procedencia.



Figura 17: Porcentuales totales de alumnos por tiempo destinado al estudio.

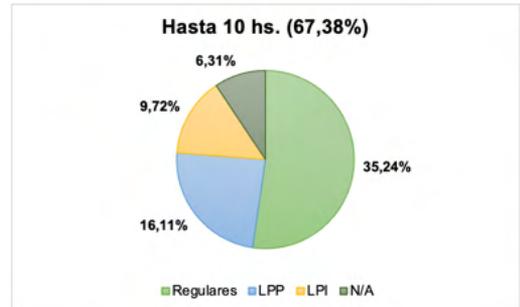


Figura 18: Alumnos que dedican hasta 10 hs. al estudio por resultados.

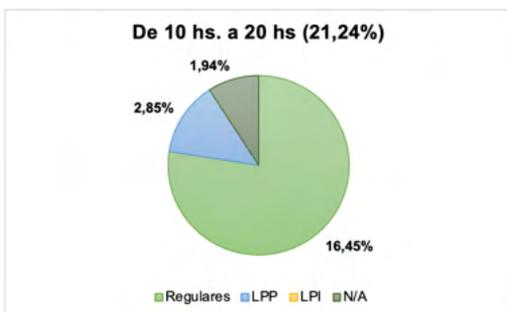


Figura 19: Alumnos que dedican entre 10 hs. y 20 hs. al estudio por resultados.

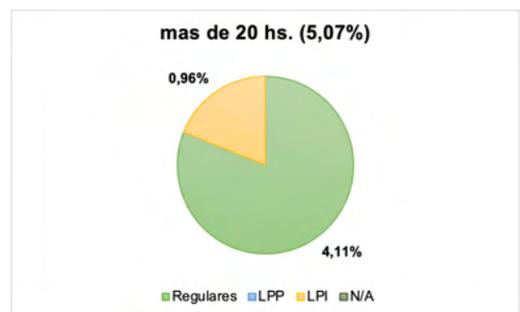


Figura 20: Alumnos que dedican más de 20 hs. al estudio por resultados.

Figura 20 referida a los alumnos que dedican más de 20 hs. semanales a sus estudios.

De la información brindada por la Figura 17 surge que tenemos una proporción aproximada de 67%, 21%, 5% y 6% de alumnos que dedican menos de 10 hs., entre 10 hs. y 20 hs. y más de 20 hs. semanales a sus estudios y ns/nc.

De las Figuras 17, 18, 19 y 20 observamos que el porcentual total de alumnos regulares que dedicaron hasta 10hs. semanales a sus estudios del 35,24%, indica que un 52,30% del total de alumnos de esa franja regularizaron la asignatura; mientras para que los alumnos que estudiaron entre 10 hs. y 20 hs. semanales y regularizaron la asignatura el porcentual total del 16,45% indica un porcentual relativo del 77,45% del total de los alumnos que estudiaron entre 10 hs. y 20 hs. semanales regularizaron la asignatura y de los valores de los alumnos que estudiaron más de 20 hs. semanales y regularizaron la asignatura se tiene que el porcentual total del 4,11% que indica un porcentual relativo del 81,07% del total de los alumnos de esa franja regularizaron la asignatura, lo que arroja una diferencia de porcentuales relativos de 28,77% entre los alumnos que estudiaron hasta 10 hs. semanales y más de 20 hs. semanales.

Para los alumnos Libres por parciales se registraron los siguientes valores, para alumnos que dedicaron hasta 10 hs. semanales el porcentual total de alumnos resulta un 16,11%, lo que indica un 23,91% del total de alumnos que estudiaron hasta 10 hs. y quedaron libres por parciales en la asignatura; en el caso de alumnos que estudiaron entre 10 hs. y 20 hs. semanales el 2,85%, indica un 13,42% del total de alumnos que estudian entre 10 hs. y 20 hs. semanales quedaron libres por parciales en la asignatura y para los alumnos que estudian más de 20 hs semanales se registró un 0% en total para esta subvariable, que resulta un 0% del total de alumnos que estudian más de 20

hs. que quedaron libres por parciales en la asignatura; resultando así una diferencia entre el mayor valor y el menor valor de 23,91% entre alumnos que estudiaron hasta 10 hs. semanales y alumnos que estudiaron más de 20 hs. semanales.

En alumnos libres por inasistencias el porcentual total de alumnos que estudiaron hasta 10 hs. semanales del 6,31%, indica un 9,36% del total de alumnos de esa franja y quedaron libres por inasistencias en la asignatura; para los alumnos que estudiaron entre 10 hs. y 20 hs. semanales el porcentual total fue 0%, que indica un 0% del total de alumnos de esa franja que quedaron libres por inasistencias en la asignatura y para los alumnos que estudiaron más de 20 hs. con un porcentual total del 0,96% que tiene una incidencia del 18,93%; resultando así una diferencia entre el mayor valor y el menor valor de 18,93%, entre alumnos que estudiaron más de 20 hs. y entre 10 y 20 hs. semanales.

Sometemos estos valores a una prueba del test del chi cuadrado.

El Chi de tabla está obtenido para un nivel de confiabilidad del 95%, encontramos que al ser el Chi calculado = 23,7023996 > Chi de tabla = 16,9189776, indica que las diferencias de valores obtenidos entre resultados observados y los esperados no son aceptables dentro del margen de lo aleatorio y *se deben a la influencia de la variable analizada.*

RESULTADOS

De los valores analizados surge que *la variable tiempo dedicado al estudio es incidente en los resultados obtenidos por los alumnos en el curso*; las variables sexo y provincia de procedencia no son incidentes. Tenemos datos de otras variables como escuela de procedencia, ocupación de los padres, situación laboral del alumno, uso de TICs de las cohortes 2018 y 2019 y de la Facultad de Ciencias Exactas y Naturales y Agrimensura

	Regulares	LPP	LPI	N/A	Totales
Hasta 10	73	33	20	13	139
10 a 20	34	6	0	4	44
más de 20	8	0	2	0	10
NS/NC	6	3	0	4	13
Totales	121	42	22	21	206

Tabla 1: Valores observados discriminados por el tiempo dedicado al estudio

	Regulares	LPP	LPI	N/A
Hasta 10	81,6456311	28,3398058	14,8446602	14,1699029
10 a 20	25,8446602	8,97087379	4,69902913	4,48543689
más de 20	5,87378641	2,03883495	1,06796117	1,01941748
NS/NC	7,63592233	2,65048544	1,38834951	1,32524272
Totales	121	42	22	21

Tabla 2: Valores esperados discriminados por el tiempo dedicado al estudio

Chi calculado	23,7023996
Chi tabla	16,9189776

Tabla 3: Valores de Chi

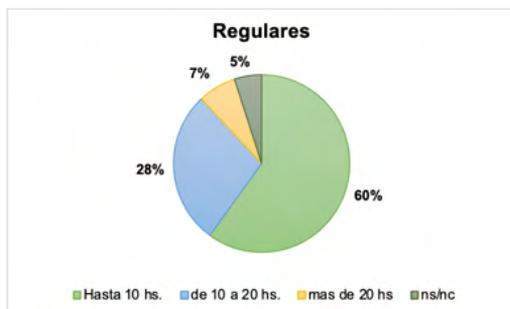


Figura 21: Alumnos Regulares por tiempo dedicado al estudio.



Figura 22: Alumnos Libres por parciales por tiempo dedicado al estudio.



Figura 23: Alumnos libres por inasistencias por tiempo dedicado al estudio.

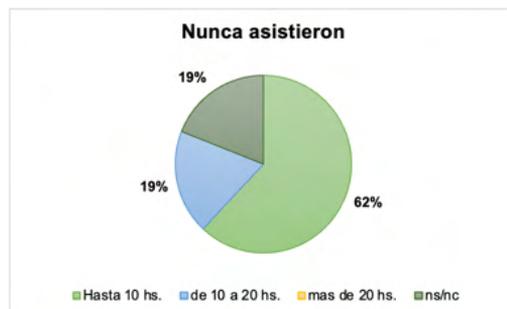


Figura 24: Alumnos que nunca asistieron por tiempo dedicado al estudio.

(FaCENA) de la Universidad Nacional del Nordeste (UNNE), que serán analizados en trabajos posteriores.

REFERENCIAS

CURTO DÍAZ, J. **Introducción al business intelligence**. UOC: Barcelona. 2010.

FORMIA, S. y LANZARINI, L. Caracterización de la deserción universitaria en la UNRN utilizando minería de datos. **Revista Iberoamericana de Tecnología en Educación y Educación en Tecnología (TE&ET)**, (11):92-98. 2013.

LA RED MARTINEZ, D.; GIOVANNINI, M.; BAEZ MOLINAS, M.; TORRE, J. y YACCUZZI, N. Academic performance problems: A predictive data mining-based model. **Academia Journal of Educational Research**; 5 (4), 61-75. 2017.

MOLINA LOPEZ, J. Y GARCIA HERRERO, J. **Técnicas de Análisis de Datos**. Madrid: Universidad Carlos III. 2006.

PEREIRA, R., ROMERO, A. y TOLEDO J. Descubrimiento de perfiles de deserción estudiantil con técnicas de minería de datos. **Vínculos**. (10) 1, 374-383 . 2013.

SALTON, G. **Automatic Text Processing: The Transformation, Analysis and Retrieval of Information by Computer**. Boston: Addison- Wesley Longman Publishing Co. 1989.