

REDISEÑO DE PROCESOS DE GESTIÓN DE LA ENSEÑANZA BASADO EN EL ANÁLISIS DE DATOS

REDESIGN OF TEACHING MANAGEMENT PROCESSES BASED ON DATA ANALYSIS

Jobany José Heredia Rico¹, Aida G. Rodríguez Hernández²

¹ Ingeniero Industrial. Profesor Adiestrado, Instituto Superior Politécnico. José Antonio Echeverría. CUJAE, Marianao, Ciudad de La Habana, Cuba.

² Dra. en Ciencias Técnicas, Profesora Titular, Profesora Consultante, Instituto Superior Politécnico José Antonio Echeverría. CUJAE, Marianao, Ciudad de La Habana, Cuba. aida@ind.cujae.edu.cu.

RESUMEN

Las instituciones de educación superior podrían beneficiarse considerablemente con la aplicación sistemática de técnicas de análisis de datos, pues esto permitiría aprender sobre el comportamiento de los procesos y de esta forma asistir en la gestión de los mismos. En el trabajo se muestran los resultados del diseño y aplicación de una metodología basada en la utilización de técnicas de análisis multivariado para obtener información relevante sobre las características de aprendizaje de los estudiantes de Ingeniería Industrial de la CUJAE. Mediante la aplicación del análisis factorial, del análisis cluster y del análisis factorial de correspondencias, se logra caracterizar el estado de aprendizaje de los estudiantes en cada curso e investigar las transiciones de uno a otro estado a lo largo del proceso de formación. Esto posibilita rediseñar la gestión de las acciones educativas para que estén más enfocadas en las diferencias individuales de los estudiantes y por tanto para que sean más efectivas.

Palabras claves: Análisis multivariado, educación superior, gestión del conocimiento, proceso de enseñanza, rediseño de procesos.

ABSTRACT

Higher learning institutions could widely benefit from regularly using data analysis since this would allow learning about processes performance and then backing up their management. The study shows the design and application of a methodology based on using multivariate analysis techniques for getting relevant information about learning characteristics of Industrial Engineering students from CUJAE. By means of using factor analysis, cluster analysis and correspondence factorial analysis, it is possible to characterize students learning condition at each academic year and investigate transitions from one condition to another throughout formative process. All this make possible the redesigning of educational actions management, with the goal that these actions be more focused on students' individual differences and hence be more effective.

Keywords: Higher education, knowledge management, multivariate analysis, process redesign, teaching process.

INTRODUCCIÓN

La educación superior cubana, y particularmente la Facultad de Ingeniería Industrial del Instituto Superior Politécnico “José Antonio Echeverría” (CUJAE), desde hace algunos años se desarrolla en un entorno con nuevas características que exigen la búsqueda de métodos para perfeccionar el trabajo en este tipo de instituciones. Algunas de estas nuevas particularidades son:

- El aumento de la matrícula. Este cambio cuantitativo ha tenido aparejado un cambio cualitativo, por cuanto ahora ingresan a la carrera, además de alumnos de muy altos rendimientos, otros con rendimiento medio y bajo, creándose una diversidad mucho mayor que en años anteriores, lo que complica el trabajo en el aula.
- Un porcentaje alto de los profesores que forman parte del claustro son jóvenes universitarios recién graduados. Este aspecto adquiere más relevancia cuando se analiza en conjunción con el anterior. Es decir, el aumento de matrícula, al provocar un cambio cuantitativo y cualitativo, demanda no sólo incrementar la cantidad, sino también la calidad de los profesores, cuestión ésta más desafiante dado el grado de experiencia del actual claustro.

Estos factores provocan la urgencia de perfeccionar la gestión del proceso docente, aprovechando las propias técnicas de ingeniería industrial para guiar de mejor manera el trabajo del profesor, brindarle informaciones más completas y oportunas sobre sus alumnos, que le faciliten orientarlos de acuerdo a sus características específicas, y controlar todo el proceso de aprendizaje teniendo en cuenta esa variabilidad. El enfoque de procesos en la formación de ingenieros industriales requiere centrarse en el alumno, y estudiar cómo se produce la transformación de los conocimientos, habilidades y otras características de entrada, en características de salida, y convertir los resultados de esos estudios en conocimientos útiles para dirigir el proceso de aprendizaje en condiciones de masividad.

Las universidades deben ser modelos del nuevo paradigma de la organización que aprende (learning organization), que plantea a las instituciones las exigencias de aprender con la experiencia, de conservar el conocimiento, y de saberlo aplicar, como requisitos imprescindibles para el éxito. En las universidades debe primar un estilo de dirección que se proponga metas, asigne recursos y controle resultados a partir de un despliegue de inteligencia organizacional, que gestione un aprendizaje organizacional basado en la utilización de las técnicas modernas de análisis de datos para que todos aprendan del comportamiento de los procesos que dirigen, puedan prever sus reacciones y dominarlos.

Esta situación, unida a la ventaja que representa para estos fines el sistema de información con que cuenta la Facultad donde se registra un gran volumen de datos viables de ser analizados, promueve el despliegue de una estrategia de análisis de datos para el descubrimiento y la gestión del conocimiento asociado al proceso de enseñanza, que contribuya a transformar los trabajos de dirección y a sugerir nuevos enfoques y acciones de los actores directos del proceso, para mejorar la productividad y eficacia de profesores y estudiantes. Este trabajo va encaminado a aprovechar con esos fines la oportunidad que representa la exploración del gran conjunto de datos referentes al proceso docente, que semestre tras semestre se acumulan en la Facultad de Ingeniería Industrial, considerando la efectividad mostrada por el despliegue de este tipo de estrategias para la obtención de conocimientos útiles como apoyo a las decisiones (Fayyad *et al.*, 1996) (Luan, 2002).

Precisamente los objetivos del trabajo son mostrar los resultados alcanzados al diseñar y aplicar una metodología para la gestión de la enseñanza basada en el análisis de datos y en la posterior obtención de conocimientos, así como la utilidad de estos resultados para lograr mayor efectividad en las acciones educativas.

MATERIALES Y MÉTODOS

Para rediseñar el proceso sobre la base del análisis de datos, se empleó una metodología adaptada a partir de la metodología genérica hecha por Davenport & Short (1990) enfocada en el rediseño de los procesos de negocio a través del uso de las tecnologías de la información y las comunicaciones (TIC). En la adaptación se consideran las características de los problemas a resolver en la enseñanza de la ingeniería industrial en la CUJAE. La metodología se estructuró en las siguientes cinco etapas:

- 1) Caracterización del proceso a mejorar.
- 2) Visión estratégica y objetivos del proyecto de rediseño.
- 3) Análisis de las mejoras que facilitan las tecnologías informativas.
- 4) Análisis de datos (descubrimiento del conocimiento).
- 5) Rediseño del proceso.

En la concepción de los aspectos incluidos dentro de cada una de estas etapas se tuvo en cuenta la propuesta realizada por Davenport (2000), referente a diversos factores críticos a considerar al desarrollar proyectos para convertir datos de procesos en conocimiento. En la actualidad los mayores esfuerzos en estos proyectos se concentran en la adquisición de una tecnología con la mayor capacidad para recolectar y almacenar grandes volúmenes de datos. No obstante, para obtener la deseada "información útil" que permita ejecutar decisiones informadas, deben considerarse, indefectiblemente, cuestiones vinculadas a las actitudes y aptitudes de los humanos para analizar e interpretar los datos (Davenport, 2000). Dentro de estas cuestiones destacan:

- La proyección estratégica de la organización donde se ejecuta el proyecto.
- Las habilidades y experiencias de las personas para el análisis de datos.
- La cultura y estructura organizacional necesaria para poder construir una "capacidad analítica" dentro de la organización.
- Aspectos de la tecnología relacionados con la comunicación y el análisis de datos.

Para la obtención de los resultados en la etapa de análisis de datos, las fuentes de datos utilizadas fueron, el sistema de información con que cuenta la Facultad (SIGENU: Sistema de Gestión de la Nueva Universidad) y los registros que se conservan en la secretaría docente. El descubrimiento de conocimiento se logra a partir de la aplicación de técnicas de análisis multivariado, específicamente, análisis factorial (AF), análisis de clusters jerárquicos y análisis factorial de correspondencias (AFC).

El análisis factorial fue realizado a partir de la matriz de correlaciones de Spearman (y no de Pearson, porque las variables originales, en este caso las calificaciones de los estudiantes en las asignaturas, tienen escala ordinal). Para obtener las correlaciones de Spearman se siguió la recomendación de la literatura que plantea que para calcular este tipo de correlaciones, las variables con escala ordinal deben ranquearse, y con las variables ranqueadas proceder al cálculo del coeficiente de correlación con lo cual se obtendrá la correlación de Spearman (Manisera *et al.*, 2010). Finalmente con las calificaciones ranqueadas es que se realiza el análisis factorial. Es recomendado efectuar este tipo de transformaciones siempre que se realice el AF con datos ordinales (Manisera *et al.*, 2010). Como criterio técnico para determinar las componentes a conservar, se seleccionaron las que tuvieron autovalores mayores que 1, hasta completar el menor conjunto de componentes que logró representar un 60% de la varianza total, por ser una aplicación de ciencias sociales (Hair *et al.*, 1999).

Después de realizar el análisis factorial, las componentes que fueron seleccionadas (según los criterios comentados) para representar el comportamiento de los estudiantes, fueron utilizadas como variables para realizar el análisis cluster. Con este análisis se desea obtener agrupaciones

de alumnos con resultados académicos similares. Se utilizó la técnica de clusters jerárquicos, habiéndose experimentado con distancias euclidianas y distancias euclidianas cuadradas, y enlaces promedio, por centroides o de Ward. Para seleccionar la cantidad óptima de clusters se consideraron dos criterios. En primer lugar asegurar un nivel de similaridad mínima en la solución del 80 %. También el análisis secuencial de la solución para detectar momentos en los cuales se unían elementos que provocaban un descenso brusco en la similaridad y que por lo tanto suponían un buen punto de corte.

Con la finalidad de detectar tendencias en el tránsito de los estudiantes de un curso a otro, o sea, de posibles asociaciones entre la pertenencia a un cluster en primer año, y la pertenencia a otro cluster en segundo, se experimentó la utilidad del análisis factorial de correspondencias. Para esto primeramente se obtuvieron, según el método propuesto, clusters de estudiantes con los resultados docentes de segundo año (curso 2008-2009)¹. Posteriormente se realizó un análisis de correspondencia simple considerando como variables categóricas “el cluster de pertenencia del estudiante al final de primer año” (curso 2007-2008) y “el cluster de pertenencia del estudiante al final de segundo año” (curso 2008-2009). Con este análisis se pretenden investigar las posibles relaciones entre los perfiles de estudiantes de ambas etapas.

El único requisito para la validez del AF es que las variables utilizadas sean numéricas (Jhonson & Wichern, 1998) (Manisera *et al.*, 2010). Como para realizar este análisis las variables ordinales (calificaciones originales) fueron numerizadas a través del uso de las calificaciones ranqueadas (Manisera *et al.*, 2010), se garantiza el cumplimiento de este requisito. La obtención de clusters jerárquicos sólo tiene como requisito que se utilicen variables numéricas para formar los grupos (Jhonson & Wichern, 1998), cuestión ésta que se cumple pues este análisis se realiza utilizando como variables las componentes resultantes del AF, que son de hecho variables numéricas (Jhonson & Wichern, 1998). El AFC sólo debe realizarse para estudiar la relación entre variable categóricas evaluadas sobre un número n de elementos (Jhonson & Wichern, 1998) (Hair *et al.*, 1999). Esta condición se verifica en el estudio ya que el AFC se aplica para conocer la relación entre dos variables categóricas que se evalúan para cada estudiante: el cluster de pertenencia del estudiante al final de primer año y el cluster de pertenencia del estudiante al final de segundo año.

Los análisis fueron aplicados a las calificaciones de 150 estudiantes de primer año del plan de estudio C´ (curso 2006-2007) y a 140 estudiantes del plan D durante su primer y segundo año (cursos 2007-2008 y 2008-2009)².

RESULTADOS

Caracterización del proceso a mejorar

La mejora de la gestión de un proceso está precedida por el conocimiento de las principales características de éste. En este caso una caracterización del proceso de enseñanza- aprendizaje es fundamental para comprender sus necesidades de gestión y para encontrar áreas de mejora que puedan derivarse del descubrimiento de conocimiento en los datos.

El proceso tiene que ser visto desde la perspectiva del alumno, cuya educación constituye el objetivo, y que es a la vez un participante decisivo en el proceso. La persona encargada de lograr que las metas del proceso se cumplan, es el vicedecano docente, apoyado por los jefes de año, que actúan como responsables en cada una de las etapas. La estructura organizativa

1 Debe notarse que estos son los mismos estudiantes de primer año en el curso 2007-2008.

2 Del curso 2006-2007 al curso 2007-2008 hubo un cambio de plan de estudios que implicó lógicamente cambios en las asignaturas que se imparten en la carrera.

no puede estar orientada al proceso, porque los participantes trabajan en varios procesos simultáneamente: un mismo profesor trabaja en procesos de formación en más de una carrera, y además en proyectos de investigación, en maestrías, entre otras funciones. El proceso es interfuncional, porque en él participan profesores provenientes de los cuatro departamentos funcionales de la facultad; pero también es interorganizacional por la participación de profesores de casi todas las facultades del Instituto. Se vincula como varios procesos, destacando los de investigación y postgrado.

La responsabilidad por el resultado es eminentemente colectiva; las acciones de unos interactúan con las de otros, incrementando o disminuyendo su efecto. Es un proceso donde priman las decisiones no estructuradas, basadas en el conocimiento de los actores, incluyendo al alumno, que toma sus propias decisiones sobre la atención que le va a dar a las orientaciones recibidas. Los resultados de cada acción de los profesores dependen grandemente de las características propias del alumno, que son muy variadas, de sus motivaciones internas y de sus habilidades para aprender; en fin, de las posibilidades de desarrollo alcanzado por la persona en el momento del aprendizaje.

De forma clásica, la coordinación de las acciones educativas en este proceso se basa en:

- Los objetivos y tiempos asignados en el plan de estudios.
- La asignación de horarios y lugares de clase que hace la facultad.
- La aprobación y distribución cronológica de evaluaciones.
- Las reuniones del colectivo de año con la participación de los profesores y los dirigentes estudiantiles de los grupos del año.
- El proyecto educativo que se elabora para cada año.

La identificación de las principales acciones que se ejecutan en el proceso es un punto de partida para saber qué conocimiento pudiera influir en la forma en que estas acciones se ejecutan.

Visión estratégica y objetivos del proyecto de rediseño

Un factor clave para el éxito del programa de mejora basado en los datos es su inclusión dentro de un marco estratégico en la organización (Davenport, 2000). Es por esto que se recomienda la definición de los objetivos estratégicos del proceso, y en función de estos, la definición de los objetivos del programa de mejora. Es necesario que los objetivos que se persiguen con el rediseño tributen a resolver problemas importantes en la organización o al menos que ayuden a que ésta consiga ventajas o mejoras estratégicas. Si se logra la articulación de ambos objetivos, se incrementará el apoyo y el compromiso con el programa de rediseño.

Como parte del desarrollo del Sistema de Gestión de la Calidad en la Facultad de Ingeniería Industrial fue definido en la política de calidad que para el logro de la mejora continua como concepto guía en el accionar de la institución, debían lograrse los siguientes objetivos en relación al proceso docente:

1. Lograr la satisfacción de los requisitos, necesidades y expectativas de nuestros alumnos en su formación integral.
2. Desarrollar la capacidad de la Facultad de organización y desempeño de los recursos humanos de forma multidisciplinaria, cooperativa, coordinada con alianzas estratégicas entre las especialidades de la Facultad para lograr la mejora continua de la calidad de nuestros servicios.

Los objetivos del programa de mejora son:

1. Alcanzar niveles satisfactorios de calidad con los recursos disponibles. Significa que cada estudiante rinda el máximo posible para él, que logre los mejores resultados de que sea

capaz, en cuanto a objetivos instructivos y educativos. Para ello es importante que quienes lo guían tengan una evaluación razonable de sus posibilidades, y le orienten acerca de cómo aprovecharlas y estimularlas.

2. Mantener costos sociales asequibles. Debe ser una consecuencia del objetivo anterior, al mejorar no sólo la eficiencia vertical y la horizontal, sino al contribuir a que inclusive aquellos que deban salir del sistema, lo hagan con la mejor preparación posible para su participación en la sociedad como trabajadores, cualquiera que sea la profesión que ejerzan.
3. Aprovechar las TIC como agentes de aprendizaje, es decir, que actúen como mediadoras del conocimiento para el profesor al tomar sus decisiones y que también actúen como mediación para el alumno: que el ambiente desarrollado de aprovechamiento de las tecnologías para la creación de conocimiento sobre el proceso docente le proporcione a los alumnos cultura y oportunidades de participar a través de acciones de consulta, desarrollo y mantenimiento del sistema, para apoyar su formación profesional en técnicas de dirección.
4. Contribuir al desarrollo de la inteligencia organizacional en la Facultad mediante la concepción de un proceso de toma de decisiones que tenga como base la utilización de los datos que actualmente se conservan sobre el proceso, más otros que deben recopilarse, y que permita la definición de modelos del proceso que ayuden a evaluar integralmente los resultados y el efecto que tienen los recursos empleados sobre el aprendizaje del alumno.

Es evidente la contribución que los tres primeros objetivos del programa de mejora hacen al primer objetivo que tiene la facultad para el logro de la mejora continua. Así mismo se evidencia relación entre el tercer y cuarto objetivo de mejora con la segunda meta establecida en la política de calidad.

Análisis de las mejoras que facilitan las tecnologías informativas

Son indiscutibles las ventajas de las TIC en el registro, almacenamiento y procesamiento de datos de cualquier proceso. No obstante, se necesita además considerar las posibilidades que éstas brindan en otros aspectos, como por ejemplo, sus posibilidades de potenciar el análisis derivado de los datos procesados, su capacidad para perfeccionar la comunicación entre los diferentes agentes participantes en la creación del conocimiento y su contribución a que utilicen mejor el conocimiento (Davenport, 2000).

Las computadoras se utilizan como apoyo a la dirección del proceso docente, sobre todo en la informatización de los procesos desarrollados en las secretarías docentes, permitiendo manejar la información de un alumnado que se ha multiplicado varias veces, haciendo el mismo trabajo más rápido y con mayor calidad de la información. El rediseño adiciona la aplicación de análisis estadísticos complejos del proceso; además, por las características del proceso expuestas, es imprescindible evaluar los beneficios de las TIC para el intercambio de informaciones entre participantes de tan variadas procedencias y para optimizar la eficiencia en el trabajo colaborativo, al compartir y construir conocimientos sobre el proceso y los factores que lo influyen.

Análisis de datos (Descubrimiento del conocimiento)

En el desarrollo del proceso docente continuamente se generan datos sobre resultados y las condiciones que dan lugar a los mismos. Por lo tanto, siguiendo uno de los enfoques clásicos de la Ingeniería Industrial, es fundamental para conocer la mejor manera en que los recursos deben interactuar con los estudiantes, analizar estos datos y conseguir información sobre las relaciones entre las variables del proceso que permita guiar oportunamente las decisiones relacionadas con los métodos de trabajo a utilizar.

Se ha utilizado el concepto, propio de las aplicaciones industriales de este siglo, de que la comprensión científica de un proceso se alcanza cuando las fuentes críticas de variabilidad están

identificadas y explicadas, la variabilidad está controlada por el proceso y los atributos de calidad de los resultados son predecibles (FDA, 2004). Estos análisis tienen como objetivo modelar el proceso de formación como un proceso estocástico con espacio de estados finito, de modo que en cada momento se caracteriza de forma integral el estado del aprendizaje demostrado por un alumno, y se investigan las transiciones de uno a otro estado a lo largo de todo el proceso de formación.

Evaluación del rendimiento académico

En este caso el objetivo es obtener una evaluación sintética del rendimiento docente del estudiante en el curso, con la menor pérdida posible de información cualitativa. Las técnicas para sintetizar la calidad del resultado de un alumno dependen de las formas de evaluación adoptadas. Los reglamentos cubanos para los cursos regulares contemplan la matrícula de ciertas materias predeterminadas para cada año académico, lo cual en otros contextos ha llevado a sintetizar los resultados a través de la calificación promedio de todas las materias y del número de materias para segunda oportunidad (Moral de la Rubia, 2006), con total pérdida de la información acerca de la variabilidad de su aprendizaje en las diferentes materias.

Es por eso que para la evaluación del rendimiento se propone la obtención de "clusters de estudiantes" tomando como referencia sus calificaciones en el curso, de forma que se logren identificar las características propias de cada estudiante mediante la caracterización del cluster al que pertenece. La evaluación del rendimiento académico de un alumno a través de las características de su cluster de pertenencia supera las tradicionalmente utilizadas que se basan en un promedio de calificaciones (Chow, 2004) (Moral de la Rubia, 2006), ya que con éstas se igualan resultados de estudiantes con calificaciones muy diferentes según el tipo de materia, dificultando encontrar factores explicativos para valores que esconden diferencias cualitativas. Para obtener los clusters se considera inadecuado utilizar como variables las notas de las asignaturas cursadas, entre otras cosas porque al ser considerable el número de asignaturas, sería muy engorroso asignarle un significado a los clusters obtenidos. Dada esta dificultad y además tratando de aprovechar la correlación existente entre las diferentes notas, previo al agrupamiento se propone la aplicación de un Análisis Factorial para sintetizar los indicadores que reflejan los resultados docentes. Al aplicar este análisis se pretenden obtener transformaciones que sinteticen los resultados de cada estudiante mediante un mínimo número de variables independientes, en este caso el número de componentes seleccionadas.

El AF ha sido aplicado a las calificaciones de los estudiantes de primer año durante dos cursos (curso 2006-2007 y curso 2007-2008) y a estudiantes de segundo año durante el curso 2008-2009. A las componentes seleccionadas como significativas (cuatro en todos los cursos) se les ha asignado un significado de acuerdo con las asignaturas que las caracterizan, es decir, de acuerdo con las materias con las que más relación presentan. A continuación se muestran a manera de ejemplo los resultados obtenidos al aplicar este análisis en el caso de primer año curso 2007-2008. En la figura 1 y 2 se muestra de forma gráfica el valor de los pesos de cada variable en cada componente para este ejemplo. Estos pesos representan de manera aproximada la correlación entre las variables y las componentes (Jhonson & Wichern, 1998).

Relación entre variables y componentes

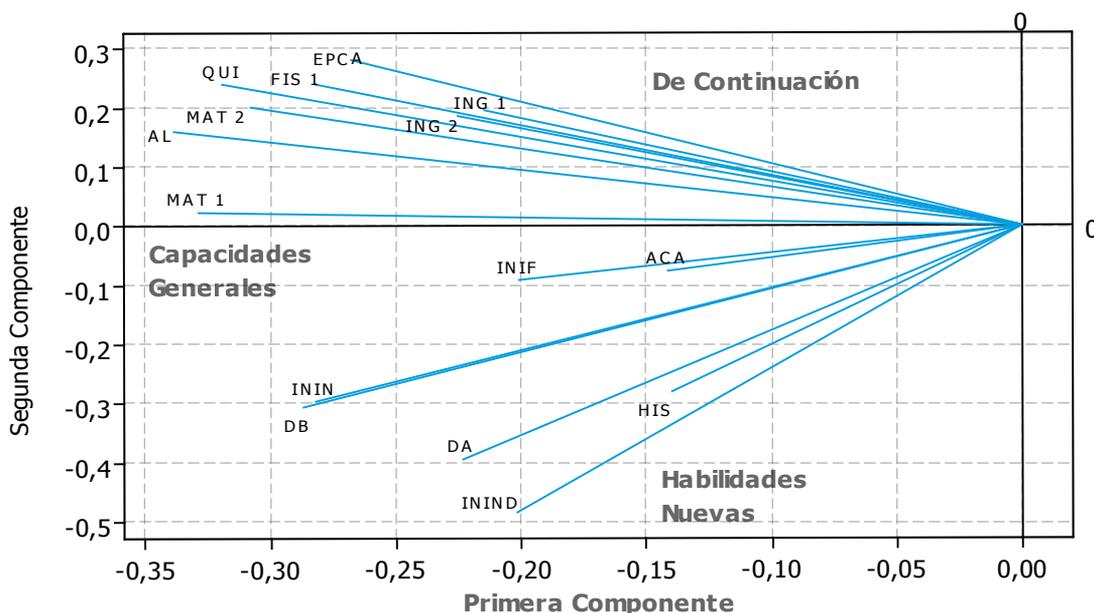


Figura 1. Primer año. Curso 07-08. Primera y segunda componentes.

En este año, al igual que en todos los restantes, todas las asignaturas cargan en el mismo sentido en la primera componente, sintetizando como lo más característico del patrón de variabilidad, que las calificaciones de los estudiantes en todas las asignaturas tienden a variar de manera conjunta. Esto significa que, frecuentemente, un estudiante con altas (bajas) notas en una asignatura, también obtendrá altas (bajas) notas en las demás. Comportamientos como éste han sido hallados en entornos diferentes en otros estudios de rendimiento estudiantil (Johnson, 1998), y se ha asociado a un factor denominado de “capacidad general de aprendizaje del estudiante”.

En todos los años, la segunda, tercera y cuarta componentes, reflejaron la existencia de subconjuntos de estudiantes con resultados opuestos en determinados grupos de asignaturas. Puede visualizarse en el gráfico 1 que en el caso de ejemplo, la segunda componente separa las asignaturas “de continuación” (que tienen un nexo fuerte con estudios anteriores) de las asignaturas que los enfrentan a conocimientos de nueva naturaleza (con perfil más técnico). Contraponiendo principalmente la Economía Política (EPCA), la Química (QUI), la Física 1 (FIS 1), la Matemática 2 (MAT 2) e Inglés 1 y 2 (ING 1 y 2), con la Introducción a la Ingeniería Industrial (ININD), Dibujo Básico y Aplicado (DB y DA), Introducción a la Ingeniería (ININ) y la Historia (HIS). Puede sorprender que Historia quede en este grupo, pero en realidad en el nuevo plan de estudio se concibe que en esta asignatura se profundice en el estudio de la Historia de localidades específicas, y sus exigencias requieren de habilidades de naturaleza bastante nueva para los estudiantes.

Como se muestra en la figura 2, la tercera componente ponen en contraposición los resultados de las asignaturas de Inglés (ING 1 y 2), Introducción a la Ingeniería (ININ) e Historia (HIS), con las calificaciones de las materias de Matemática 1 y 2 (MAT 1 y 2), Dibujo Aplicado (DA) y Apreciación de la Cultura y el Arte (ACA). En la cuarta componente el contraste evidente se produce entre los resultados de las asignaturas de Introducción a la Ingeniería Informática (INIF) y Apreciación de la Cultura y el Arte. No ha sido evidente caracterizar, como se hizo en la segunda componente, cada uno de los subgrupos de asignaturas que se separan en estas componentes,

no obstante esto no disminuye el valor de los resultados obtenidos, pues lo esencial es conocer que en ese año hubo subconjuntos de alumnos con resultados opuestos en estos grupos de materias.

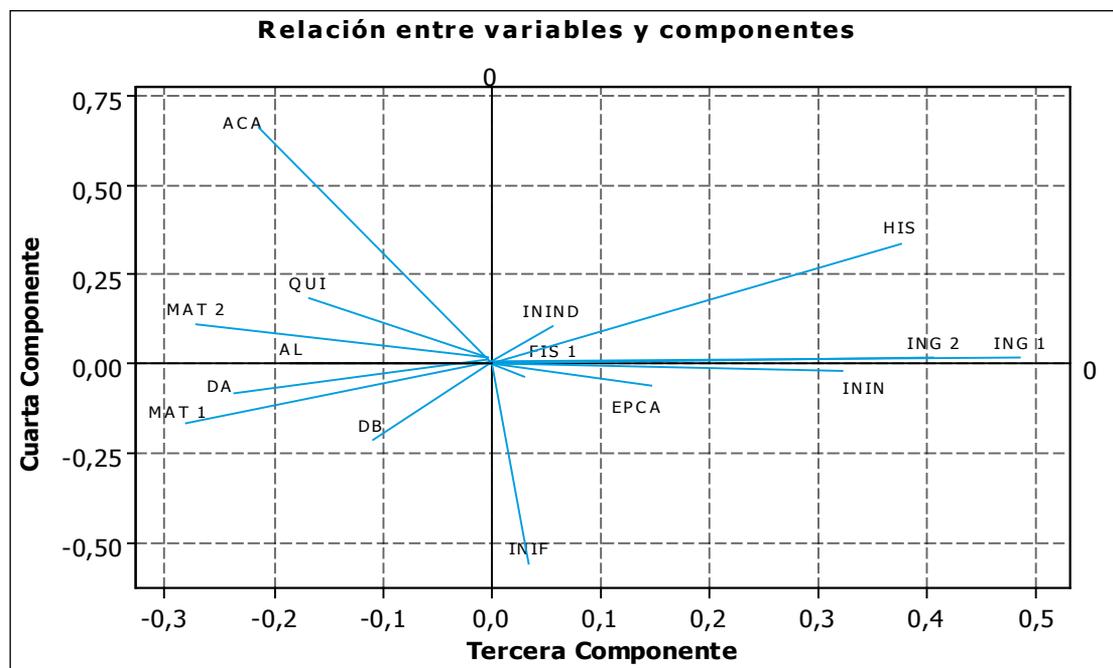


Figura 2. Primer año. Curso 07-08. Tercera y cuarta componentes.

Vale destacar que con los datos de los estudiantes de primer año curso 2006-2007 (Plan C') se obtuvieron resultados diferentes en la segunda y tercera componentes, pues la segunda componente separó asignaturas que requieren del estudiante más capacidad de abstracción (Álgebra, Química y Matemáticas) de aquellas que requieren más habilidades para la aplicación práctica del conocimiento (Fundamentos de Ingeniería II, Problemas Prácticos de Ingeniería Industrial e Inglés 1 y 2). Por su parte la tercera componente expresó una contraposición entre los idiomas, las matemáticas, Física, Química y Filosofía y Sociedad de un lado, y los dibujos, Computación, Fundamentos de Ingeniería y Problemas Prácticos de Ingeniería Industrial del otro, estableciendo una diferencia en el aprovechamiento de los alumnos entre las materias más básicas y las más profesionalizantes.

Después de aplicar el AF, las 4 componentes seleccionadas en cada uno de los casos fueron utilizadas para obtener agrupaciones de alumnos con resultados académicos similares mediante la técnica de clusters jerárquicos. Cada uno de los clusters representa un estado diferente de aprendizaje, cuya interpretación está dada por su ubicación en los ejes factoriales (Rodríguez & Haedo, 2008). Así por ejemplo, los clusters formados por alumnos con un alto valor en la primera componente y valores pequeños en las demás componentes, corresponden a estudiantes con un excelente aprendizaje en todas las materias del año. Éste es, por supuesto, un subgrupo con aprendizaje superior a todos los demás. Pero no entre todos los clusters es procedente establecer una relación de tipo ordinal, ya que generalmente se encuentran varios clusters de alumnos de buen rendimiento, pero con rendimientos superiores principalmente en ciertas materias más que en otras, o varios grupos con rendimiento medio, pero desiguales en cuanto a las materias en que mejor avanzan.

Por ejemplo, como se muestra en el dendrograma de la figura 3, obtenido al aplicar el análisis cluster con las componentes seleccionadas en primer año curso 2007-2008, se encontró que la

clasificación lograda con enlace de Ward es más fácil de interpretar y aprovechar, produciendo dos grupos de alto rendimiento, cuatro grupos de rendimiento medio y dos grupos de bajo rendimiento. A su vez, las diferencias dentro de cada categoría (alto, medio y bajo rendimiento) están dadas por el tipo de materias en las que lograron mejor aprendizaje. De forma que en el grupo de alto rendimiento un primer subgrupo resultó de alto rendimiento en todas las materias y el otro presenta sus mejores resultados en las asignaturas que requieren nuevas habilidades. De modo similar ocurrió con los clusters de alumnos de rendimiento medio, pues un grupo representa alumnos con sus mayores aptitudes en asignaturas de continuación; otro, estudiantes mejores en las materias de Arte e Historia, etc. Esto da excelentes oportunidades para orientar el trabajo de los profesores (Rodríguez & Haedo, 2008).

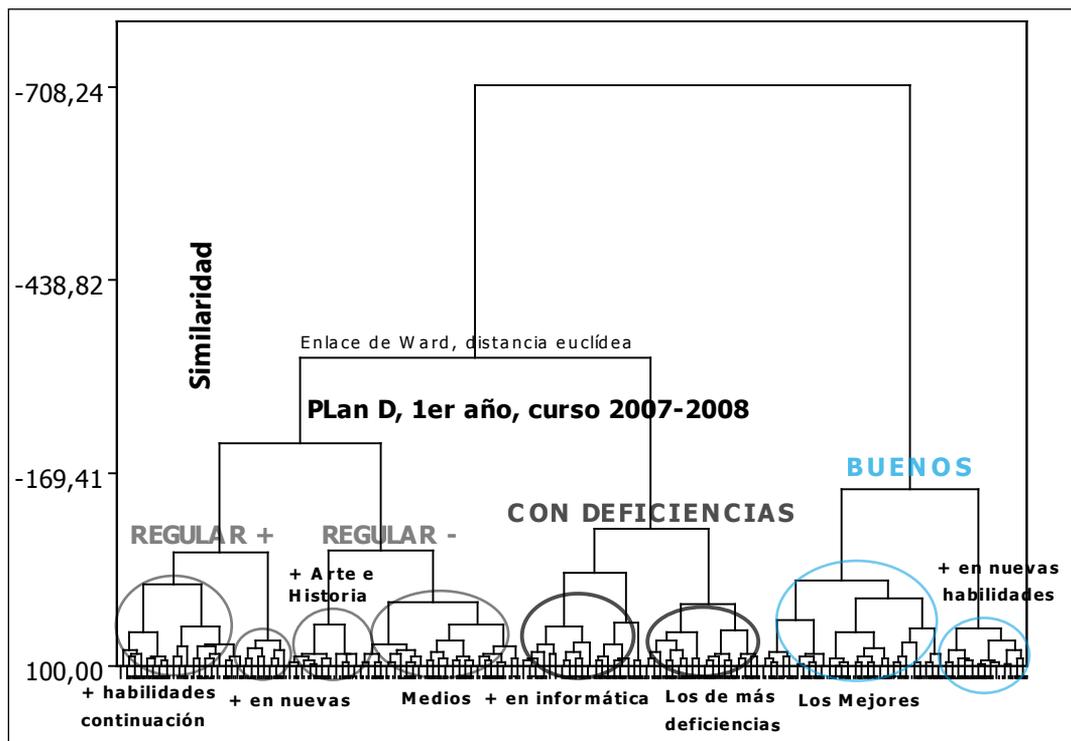


Figura 3. Agrupamiento Primer Año. Curso 07-08. Materias en que son mejores (+).

Relación entre el rendimiento en primer y segundo año

Se realizó el análisis de correspondencias y se seleccionaron dos componentes para representar los perfiles filas y columnas, lo que significa que se representarán los vectores asociados a estos perfiles en el plano que maximice la inercia de los puntos proyectados, con el fin de mantener la diferencia entre estos perfiles lo más similar posible a la diferencia que existe entre los mismos en los espacios originales. Con dos dimensiones se logra explicar el 72,48% de la inercia total. En la figura 4, es el denominado “ploteo simétrico de filas y columnas”, en el cual, la cercanía entre las categorías de una variable con las categorías de la otra, implica que son categorías que presentan mayor asociación. Cada una de las categorías de las dos variables categóricas utilizadas en el AFC, es decir, cada uno de los clusters formados en primer y segundo año, fueron etiquetados con nombres alegóricos a las características de cada cluster. Las etiquetas (siglas) que son mostradas en la figura 4, se obtuvieron de acuerdo a como se observa en la tabla 1:

Tabla 1. Significado de las etiquetas usadas en el AFC

	Primer año (2007-2008)	Segundo año (2008-2009)
1: Etiqueta para clusters de alumnos con alto rendimiento	1: Grupo con los mejores alumnos 1 N: Grupo de alto rendimiento pero con más habilidades en las asignaturas de nueva naturaleza	1: Grupo con los mejores alumnos 1-IG: Grupo de alto rendimiento con deficiencias en el Inglés 1-DF: Grupo de alto rendimiento con deficiencias en las asignaturas de aplicación militar
2: Etiqueta para clusters de alumnos con rendimiento medio	2: Grupo de rendimiento medio sin ninguna particularidad 2 N: Grupo de rendimiento medio con más habilidades en las asignaturas de nueva naturaleza 2+CO: Grupo de rendimiento medio con más habilidades en las asignaturas de continuación 2+HACA: Grupo de rendimiento medio con más habilidades en las asignaturas de Historia y Arte	2: Grupo de rendimiento medio sin ninguna particularidad 2-M4: Grupo de rendimiento medio con menos habilidades en la asignatura Matemática 4 2-PRO: Grupo de rendimiento medio con menos habilidades en las asignaturas de programación en computadoras
3: Etiqueta para clusters de alumnos con bajo rendimiento	3: Grupo con más deficiencias 3IF: Grupo con deficiencias pero con habilidades en Informática	3: Grupo con más deficiencias 3+: Grupo con deficiencias pero muy cercano a los medios 3+MEP: Grupo con deficiencias pero con buenos resultados en Estadística

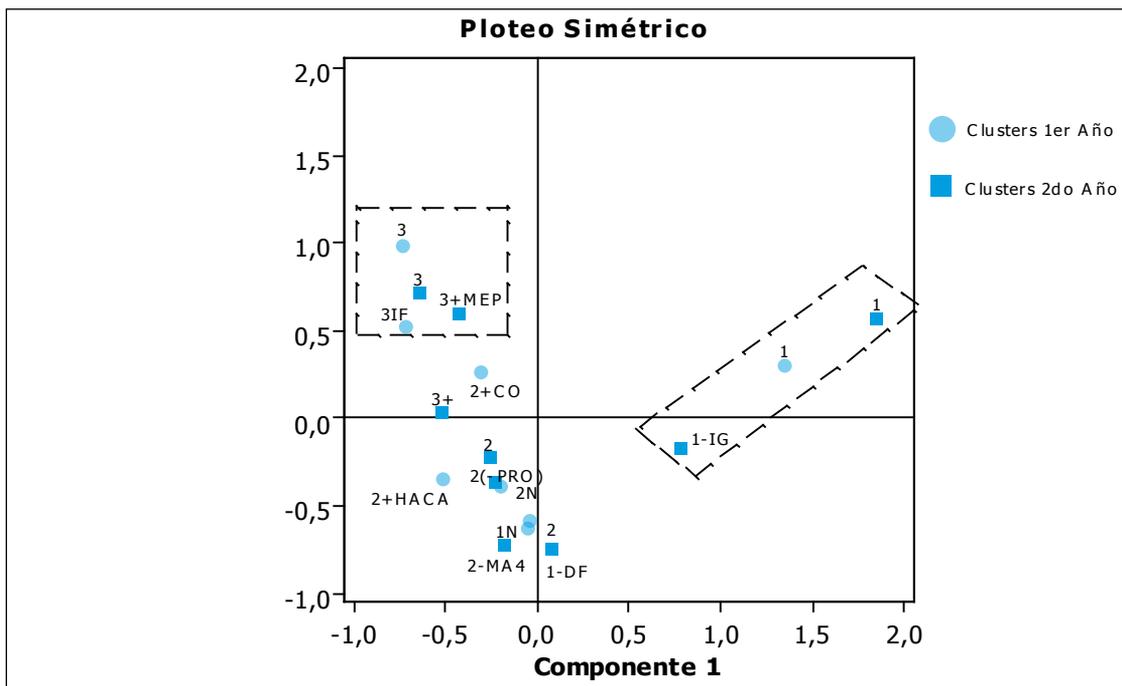


Figura 4. Correspondencia entre clusters de primero y segundo año.

En primer lugar se observa que el cluster etiquetado como 1 en primer año (el de mejores estudiantes) está asociado fundamentalmente con dos clusters de segundo año, el 1 y el 1-IG, ambos formados también por estudiantes de alto rendimiento. El otro resultado que resalta es que los grupos etiquetados como 3 y 3IF en primer año (grupos de bajo rendimiento con resultados favorables sólo en algunas asignaturas), están asociados principalmente a los clusters marcados como 3 y 3+MEP en segundo año, los cuales están formados igualmente por estudiantes de bajo rendimiento. Por lo tanto este análisis puso de relieve que las tendencias existen en el

sentido esperado, es decir, los mejores al finalizar una etapa, tienden a ser los mejores al finalizar la siguiente. Similar tendencia se manifiesta entre los alumnos con dificultades. No obstante también se detectó en la tabla de contingencia resultante de este análisis, que con carácter no excepcional subconjuntos de alumnos transitan a un estado bastante diferente del inicial (algunos alumnos que fueron muy buenos en todas las materias en primer año, bajaron notablemente su rendimiento en segundo año, y también viceversa), en cuyas causas es preciso profundizar el análisis, para descubrir factores que deben ser reforzados, si llevan a un estado definitivamente superior al esperado, o que deben ser contrarrestados, cuando llevan a estados inclusive muy inferiores al inicial.

Rediseño de la gestión del proceso de formación

La esencia del proyecto de rediseño es la gestión del conocimiento proporcionado por el análisis de datos, para influir en la eficiencia y eficacia del proceso mediante la organización de una producción colectiva de conocimientos. Este objetivo necesita una estructura organizativa que dé soporte a decisiones analíticas basadas en datos (Davenport, 2000). Para lograrlo, en el proyecto se contempla la creación de un Laboratorio de Análisis de Procesos, el cual tenga como un objetivo esencial la realización de prácticas profesionales de estudiantes de pregrado y posgrado, de modo que cada uno siempre esté guiado y supervisado por un profesional de mayor nivel científico y académico en el área de modelación de procesos. Este laboratorio se encarga no sólo del procesamiento de los datos, sino que una parte importante del entrenamiento que en el laboratorio se recibe, es a través de la participación en su interpretación y en la formalización de esas interpretaciones. La utilización del conocimiento obtenido en la ejecución del proceso de formación, está precedida por acciones que propicien que cada participante comprenda y comparta los objetivos y métodos que se emplearán, y que debe constituir una fase de capacitación y de involucramiento.

El etiquetado y descripción de las componentes, de los clusters, así como las restantes interpretaciones derivadas del análisis, deben ser realizadas bajo la dirección de los respectivos jefes de año por los especialistas del laboratorio, con el asesoramiento de los profesores más conocedores de cada año académico. Por las características de la Facultad, en todos los años se puede contar con profesores con amplia experiencia en la docencia del año, y con conocimientos matemáticos suficientes para la tarea. Estos resultados se llevan a los colectivos de año y de disciplina, a cuyos integrantes corresponde explicar y enriquecer el análisis de causas subyacentes que los modelos matemáticos sugieran, convirtiéndolo en conocimiento profesional compartido. El estado inicial de cada alumno será comunicado al inicio del período a todos los integrantes del colectivo de año en esa etapa, para que la orientación y el control del profesor estén más adecuados a las potencialidades demostradas del alumno, sin dejar por ello de estimularlo a alcanzar resultados superiores.

Debido a que todos los involucrados se desempeñan en varios procesos a la vez, no tiene sentido pasar a una estructura de procesos; el vicedecano, el coordinador del laboratorio y los jefes de año requieren desarrollar sus habilidades de persuasión, desplegando todo el atractivo del trabajo en el equipo, y minimizando sus inconvenientes. Los modelos matemáticos del proceso podrán usarse en el laboratorio para simular nuevos entornos y evaluar los resultados más probables; la confrontación posterior con la realidad permitirá el refinamiento y ajuste progresivo de los modelos y de los métodos como base para las decisiones. El laboratorio se desempeña, pues, no como una unidad administrativa de procesamiento de datos, sino como una unidad docente y de investigación, cuyo objeto de estudio es el proceso de formación.

DISCUSIÓN

A partir de los resultados obtenidos con el AF se logran caracterizar las interdependencias entre resultados de grupos de asignaturas, pues con el comportamiento manifiesto en la primera componente se conoce cuál es la relación más sistemática (general) que existe entre las calificaciones en las diferentes materias; además que la interpretación de las restantes componentes, aunque proporciona información sobre comportamientos menos sistemáticos (que ocurren en grupos menores de estudiantes) también ayudan a la mejor comprensión de estas relaciones. Los cambios en el plan de estudio también se reflejan en el AF, pues por ejemplo, en la segunda componente, las materias de Inglés, cuyas calificaciones en el Plan C se relacionaban mucho con las de Introducción a la Ingeniería Industrial y Problemas Prácticos de Ingeniería, en el Plan D pasaron a relacionarse esencialmente con las materias de Física, Química y Matemática.

Con el análisis de agrupamiento se identifican tipos de alumnos según sus competencias académicas logradas al finalizar cada etapa, evaluándose no solamente las dificultades que presentan, sino también sus potencialidades. Además este análisis permite sintetizar en una sola variable las principales características del aprendizaje del alumno, lo cual ayuda en el planteamiento de un modelo explicativo del rendimiento académico. Específicamente con la metodología se obtiene una variable categórica (el cluster de pertenecía del estudiante al final de un curso) que caracteriza el rendimiento académico del alumno. Estudios de este tipo han sido realizados en otras universidades (Chow, 2004) (Reyes *et al.*, 2007) (Bañadme *et al.*, 2007) con el fin de investigar la relación entre este rendimiento y variables nominales, ordinales o cuantitativas que expresan las características de entrada a la universidad del alumno, y de las que expresan las características de los recursos utilizados en el proceso. Actualmente se investiga la utilidad de la regresión logística multinomial para explicar la pertenencia a uno u otro de los clusters que fueron formados en primer año, a partir de variables independientes como las que fueron comentadas. Otra de las técnicas que se evaluará para ese propósito es la modelación de ecuaciones estructurales, como en el modelo obtenido en la Universidad de Quintana Roo, México para monitorear el impacto que tienen las acciones y los recursos empleados en los cursos de ingeniería, indicando cuáles procesos requieren atención prioritaria (Cuevas & Rodríguez, 2007).

Con el AFC se detectó un comportamiento sistemático en el rendimiento estudiantil al pasar de primero a segundo año, por lo que esta técnica pudiera ser una vía para predecir, a finales de primer año, el comportamiento de los estudiantes en segundo. En efecto si se determinó que la distribución condicional de las frecuencias para los clusters de segundo año no es independiente de los resultados de primero, es decir, la probabilidad de que un alumno pertenezca a determinados clusters en segundo año está influenciada por el cluster al que perteneció en primer año, pudiera utilizarse esta distribución (grupo al que pertenecerá en 2do dado el grupo al que perteneció en 1ro) para, al tener el resultado del agrupamiento de los alumnos de primer año de cierto curso, pronosticar a qué conglomerado de segundo año se espera que pertenezcan. Esto facilita la toma de medidas educativas preventivas con los estudiantes que lo requieran.

Ventajas del análisis para la gestión del proceso

Las acciones de gestión en el proceso de enseñanza son desarrolladas por diferentes actores y en diferentes niveles de decisión, y para cada uno de estos niveles de decisión el conocimiento descubierto podrá contribuir de manera distinta (Romero & Ventura, 2007). La propia impartición de las clases es una actividad que requiere gestionarse para lograr su correcta planificación, ejecución y control. La didáctica, como disciplina rectora de esta actividad, asiste en su dirección científica mediante la consideración de una serie de categorías (colectivo de autores, 2003). Según esta disciplina, el método, los medios de enseñanza y algunas de las actividades de control

del aprendizaje utilizadas en el proceso docente deben estar en función de los conocimientos, habilidades y motivaciones de los estudiantes (colectivo de autores, 2003). Precisamente la información útil sobre las características de aprendizaje de los estudiantes que se obtiene de los análisis de agrupamiento y de correspondencia, puede ayudar a que el profesor, junto con los conocimientos de didáctica, establezca los métodos, procedimientos, medios de enseñanza y de evaluación a utilizar en su clase que considere más efectivos según las características de los distintos estudiantes.

No se puede obviar el hecho de que el desarrollo de la actividad docente no abarca solamente la gestión individual del profesor para su asignatura, sino también la gestión general que realiza el colectivo de año. Este colectivo es dirigido por un profesor designado como jefe de año, y está formado por todos los profesores que imparten docencia en ese año. Su principal tarea es la confección del proyecto educativo, documento en el cual se identifica y aborda la solución de las principales necesidades educativas, grupales e individuales de los estudiantes, a partir de realizar un diagnóstico de sus particularidades (Horruitinier, 2007). Los análisis propuestos contribuyen considerablemente en este sentido ya que el colectivo de año podrá contar con una evaluación de la situación de aprendizaje de cada estudiante, e incluso predecir su rendimiento esperado en el semestre. Al integrar ambas informaciones podrán perfilarse las acciones particulares que se proponga el colectivo realizar con los diferentes estudiantes, así como la proposición de actividades, tareas y recursos de aprendizaje que favorezcan al estudiante según su perfil.

Además los resultados confirman la utilidad de estas técnicas en las decisiones relacionadas con la asignación de estudiantes a grupos de investigación que requieren determinadas habilidades, la identificación temprana de alumnos con condiciones para ser alumnos ayudantes³ y el diagnóstico de las diferencias entre cursos.

El rediseño genera diversos tipos de cambios en la gestión de los procesos, en cuanto a:

- el rol de los dirigentes en la actividad rediseñada
- la estructura organizativa
- los requerimiento de competencias
- la gestión de la infraestructura de las TIC
- la implementación del mejoramiento continuo

CONCLUSIONES

La aplicación de las técnicas propuestas en el modelo permite identificar automáticamente tipos de alumnos según las materias en que tienen sus mayores habilidades o insuficiencias, identificar las interdependencias entre los resultados de grupos de asignaturas y entre los resultados de los estudiantes en diferentes cursos. Además posibilitan sintetizar en un solo indicador el aprovechamiento de cada estudiante para en etapas posteriores determinar los vínculos de este aprovechamiento con las variables de entrada y del proceso. La utilización oportuna de todo este conocimiento, lograda a partir de la consideración de los aspectos recomendados en la metodología propuesta, brinda la posibilidad de planificar y ejecutar las acciones educativas con mayor atención en las diferencias individuales de los alumnos además de dotar la gestión del proceso de un carácter más preventivo.

No obstante el diagnóstico del aprendizaje que se realiza, está centrado en la evaluación de los conocimientos y habilidades de los estudiantes según sus calificaciones, pero realmente la caracterización del aprendizaje abarca además la valoración del desarrollo del pensamiento y

3 Alumnos que imparten clases de determinada asignatura a estudiantes de años inferiores.

la formación de valores lograda por el alumno. Por lo tanto deben conjugarse sus resultados a otras evaluaciones del proceso docente, para poder juzgar la calidad del aprendizaje.

REFERENCIAS

Chow, T. (2004). Predicting End of 1st Year Performance. Annual Forum for the Association for Institutional Research, Boston, May.

Colectivo de autores. (2003). *Preparación Pedagógica Integral para profesores universitarios*. La Habana: Félix Varela.

Cuevas, J., & Rodríguez, A. (2007). Modelación de ecuaciones estructurales para la evaluación formativa de los programas educativos. *Ingeniería Industrial*, 28, 64-70.

Davenport, T.H., & Short, J.E. (1990). The New Industrial Engineering: Information Technology and Business Process Redesign. *Sloan Management Review*, 31(4), 11-27.

Davenport, T.H. (2000). *Data to Knowledge to Results: Building an Analytic Capability*. Massachusetts: Institute for Strategic Change.

Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). From data mining to knowledge discovery in databases. *Communications of the ACM*, 39(11), 37-54.

FDA. (2004). PAT - A Framework for Innovative Pharmaceutical Development, Manufacturing, and Quality Assurance. Guidance for Industry

Hair, F., Anderson, R., Tatham, R., & Black, W. (1999). *Análisis Multivariante*. Madrid: Prentice Hall Iberia.

Horruitiner, P. (2007). El problema de la calidad, el acceso y la pertinencia. *Revista Pedagogía Universitaria*, 12(4), 92-132.

Johnson, D. (1998). *Métodos multivariados aplicados al análisis de datos*. México: International-Thomson Editores.

Luan, J. (2002). Data Mining and Knowledge Management in Higher Education -Potential Applications. Annual Forum for the Association for Institutional Research, Toronto, June.

Manisera, M., Van Der Kooij, A., & Dusseldorp, E. (2010). Identifying the Component Structure of Satisfaction Scales by Nonlinear Principal Components Analysis. *Quality Technology & Quantitative Management*, 7, 97-115.

Moral de la Rubia, J. (2006). Predicción del rendimiento académico universitario. *Perfiles Educativos*, 113, 38-63
Vandamme, J-P., Meskens, N., & Superby, J-F. (2007). Predicting Academic Performance by Data Mining Methods. *Education Economics*, 15, 405-419.

Reyes, J., Escobar, C., Duarte, J., & Ramírez, P. (2007). Una aplicación del modelo de regresión logística en la predicción del rendimiento estudiantil. *Estudios Pedagógicos*, 33 (2), 101-120.

Rodríguez, A., & Haedo, Y. (2008). Análisis multivariado del proceso de aprendizaje en el 1er año de ingeniería. 14 Convención científica de ingeniería y arquitectura, La Habana, Diciembre.

Romero, C., & Ventura, S. (2007). Educational data mining: A survey from 1995 to 2005. *Expert Systems with Applications*, 33, 135-146.

Vandamme, J., Meskens, N., & Superby, J. (2007), Predicting Academic Performance by Data Mining Models, *Education Economics*, 405-419

