



**ARTÍCULO ORIGINAL
INVESTIGACIÓN DE OPERACIONES**

Análisis multivariado del aprendizaje de Ingeniería Informática en la municipalización

Multivariate analysis of the learning of the Computer Engineering in the municipality

Yendri Monteagudo-Yllobre^I, Roberlando Hernández-Milian^{II}, Daniel Yunier González-Fernández^{III}

^I Banco de Inversiones S.A. La Habana, Cuba.

E-mail: yendri@bdi.cu

^{II} Empresa del Papel. La Habana, Cuba.

E-mail: roberlando@papyrus.minbas.cu

^{III} Empresa de Servicios y Comercialización de Equipos Automotores y Manipulación de Cargas "MONCAR". La Habana, Cuba.

E-mail: danielgf@moncar.co.cu

Recibido: 07/03/2011

Aprobado: 19/03/2012

RESUMEN

La aplicación en los procesos docentes de técnicas de análisis multivariado, brinda la posibilidad de planificar y ejecutar las acciones educativas, teniendo en cuenta la variabilidad en las características de los alumnos, y permite desarrollar estrategias, métodos y procedimientos más efectivos en la potenciación del incremento del rendimiento docente de los estudiantes. El presente trabajo posee como objetivo, mostrar los resultados que se pueden obtener con la aplicación de las técnicas anteriormente mencionadas y la utilidad de estos resultados para lograr una mayor efectividad de las acciones educativas. Mediante la aplicación del análisis de componentes principales, análisis factorial común y análisis de *cluster*, se logró caracterizar el estado de aprendizaje de los alumnos estudiados e investigar los patrones y regularidades que rigen sus comportamientos. Esto posibilita rediseñar la gestión de las acciones educativas y ejecutar aquellas que el estudiante requiere para alcanzar el mejor rendimiento académico posible.

Palabras clave: análisis multivariado, modelación, procesos, aprendizaje.

ABSTRACT

The use of multivariate analysis on teaching process is very effective to plan and execute educational actions having into account the variability of student's characteristics. At the same time, it is important to develop better strategies, methods, procedures and teaching and evaluation means to increase the educational yield of students. So the aim of the current work is show the expected results of this techniques application and its relevance to increase the educational actions effectiveness. Applying the main components, the common factorial and the cluster analysis it's possible characterize the learning stage of the students on research and investigate the patterns of their behavior. These results enable the redesign of the educational

actions management and the execution of those ones that the student truly requires to get the best academic yield as possible.

Key words: *multivariate analysis, modeling, processes, learning.*

I. INTRODUCCIÓN

La alta velocidad con que aumenta y se reproduce el conocimiento, aparejado al increíble desarrollo de la ciencia y la tecnología, conducen a que las sociedades necesiten cada vez más de conocimientos y habilidades de nivel superior, frente a lo cual una de las vías de respuesta es el aumento de las matrículas universitarias [1]. Frente a esta situación, Cuba, consciente de la importancia del mencionado fenómeno para los países en vías de desarrollo y a pesar de su difícil situación económica, en los últimos años ha aumentado considerablemente los recursos financieros destinados a la Educación Superior y el número de matrículas en esta enseñanza, fundamentalmente los vinculados a los Programas de Universalización entre los cuales se encuentran los Centros Universitarios Municipales, destinados a elevar el nivel educacional del sector laboral del país. Por tanto, resulta necesario que aparejado al elevado consumo de recursos financieros de la educación superior, se obtengan resultados de calidad que justifiquen el empleo de los mencionados recursos, lo cual se logra a partir de la aplicación de técnicas de apoyo a la gestión que aseguren la eficacia de las acciones realizadas.

Cualquier proceso de aprendizaje se da a partir de un conjunto complejo de procesos interrelacionados entre sí, los cuales deben gestionarse basándose en mediciones objetivas que reflejen hechos; mucho más la enseñanza de la Ingeniería, la cual se encuentra inseparablemente ligada al desarrollo de la ciencia y la tecnología. Sin embargo, frente a la dinámica de desarrollo actual de la ingeniería de alto nivel, esto resulta insuficiente; se necesita además lograr el acercamiento progresivo de los procesos a su nivel óptimo, de manera tal que se minimice el riesgo de no alcanzar la mejor calidad posible en los resultados. Dicho fenómeno se traduce en la urgencia de perfeccionar el proceso docente, de manera tal que se logre guiar de forma óptima el trabajo del profesor en su labor de orientación y ayuda a los estudiantes, brindándole informaciones más completas y oportunas sobre las características individuales de los alumnos, que les faciliten su labor y potencien la efectividad de la misma; y permitiéndole controlar todo el proceso de aprendizaje teniendo en cuenta la variabilidad en las características de los alumnos. Se requiere centrar el proceso docente en el alumno y estudiar cómo se produce la transformación de los elementos de entrada (conocimientos, habilidades, etc.) en elementos de salida; y transformar los resultados de estos estudios en conocimientos útiles para dirigir el proceso de aprendizaje en condiciones de masividad [2].

Para lograr dicho propósito, se emplea a la modelación matemática de procesos como uno de los medios fundamentales, partiendo 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 [3]. En tal sentido, resulta indispensable la modelación matemática de los procesos de enseñanza con el objetivo de garantizar la calidad de sus resultados (materializados en la elevación del rendimiento académico de los estudiantes) por medio de la salvaguarda de la calidad de los datos, su síntesis y organización, el establecimiento de asociaciones entre ellos y su transformación en conocimiento que sirva de soporte para la identificación de áreas de mejora y, por tanto, para la toma de decisiones, mediante algoritmos y procedimientos eficaces [4].

Las técnicas de modelación de procesos universitarios más empleadas, por lo general, son las basadas en el análisis de componentes principales, el análisis factorial de componentes principales con rotación Varimax, las ecuaciones estructurales, los análisis de *clusters* (grupos), algoritmos de minería de datos para la asociación de atributos, árboles de clasificación y regresión, redes de Kohonen y redes neuronales [5]. El empleo de una de estas técnicas depende, no solo de los objetivos a alcanzar, sino también del tipo de datos que se poseen y del grado de aporte que se pretende que proporcionen los analistas en la obtención del modelo y en la interpretación de sus resultados, ya que muchos de estos métodos permiten sintetizar en una o pocas variables la calidad del resultado del proceso (lo cual constituye una de sus ventajas fundamentales), facilitando la identificación de los problemas y la toma de decisiones.

ANÁLISIS MULTIVARIADO DEL APRENDIZAJE DE INGENIERÍA INFORMÁTICA EN LA MUNICIPALIZACIÓN

Es por tanto el objetivo del presente trabajo, mostrar los resultados que se pueden obtener con la aplicación de las técnicas anteriormente mencionadas y la utilidad de estos resultados para lograr una mayor efectividad de las acciones educativas.

Se posee como objeto de estudio el proceso de enseñanza de los Centros Universitarios Municipales, ya que éste posee características particulares dentro de la Enseñanza Superior Cubana, es el sector más joven de ésta y agrupa al mayor porcentaje de la matrícula y una parte importante de los recursos financieros destinados al mencionado sector; particularmente el desarrollado por el Centro Universitario Municipal de 10 de Octubre–Arroyo Naranjo (ubicado en la capital cubana) en la impartición de la carrera de Ingeniería Informática.

II. MÉTODOS

A. Datos

Los datos empleados en el estudio se corresponden con las calificaciones obtenidas por los estudiantes de la carrera de Ingeniería Informática del Centro Universitario Municipal de 10 de Octubre–Arroyo Naranjo, en las diferentes asignaturas correspondientes a los cursos comprendidos entre septiembre del 2002, curso en que entra en funcionamiento dicho centro, y julio del 2009. Como datos complementarios se obtuvieron todos los asociados a la matrícula del estudiante, que incluyen: datos demográficos, fuente de ingreso, estudios precedentes, entre otros; los cuales, además de servir en etapas posteriores para el establecimiento de variables predictivas, se emplean como ayuda para la verificación de resultados atípicos.

Este tipo de centro de enseñanza tiene como características la posibilidad, por parte de los estudiantes, de cursar en más de una ocasión un año de la carrera que posean suspenso y que por tanto, les impida ser promovidos a años superiores (lo cual es conocido como repitencia); así como de ser promovidos a años superiores, a pesar de poseer una asignatura suspensa en un año anterior, teniendo la obligación de cursar nuevamente esa asignatura de forma paralela a las asignaturas asociadas al año superior en el que se encuentra (lo cual es conocido como arrastre). Bajo las citadas características, la mayoría de los estudiantes que matriculan la carrera no se gradúan al terminar su ciclo correspondiente (6 años en este caso), sino que quedan diluidos en la matrícula general del centro, arrastrando asignaturas incluso de años anteriores que ya cursaron y suspendieron, siendo el tiempo promedio de realización de la carrera por estudiante superior a los 6 años. Tal es el caso, que la primera graduación efectuada en julio del 2008 fue de solamente 8 estudiantes y la segunda graduación efectuada en julio del 2009 fue de 15 estudiantes.

Los alumnos con estas características constituyen la mayoría de la matrícula del centro, por tanto, ocurre el fenómeno de que sobre los alumnos matriculados en una misma asignatura o ubicados en un mismo año, actúan diferentes conjuntos de factores, ya sea por la disminución o aumento de la carga ocasionado por la repitencia de asignaturas, por las ventajas de la experiencia acumulada al haber cursado una asignatura o por el cambio del entorno académico, entre otros muchos factores. Podría pensarse entonces que estos estudiantes no deberían ser tenidos en cuenta en el análisis, puesto que los diferentes conjuntos de factores que sobre ellos actúan, de por sí los discrimina, limitando la posibilidad de descubrir verdaderos factores latentes y patrones de comportamiento que los identifiquen. Sin embargo, estos estudiantes constituyen la mayoría de la matrícula y el mencionado comportamiento caracteriza a este tipo de enseñanza, por lo que no tiene sentido excluirlos si lo que se desea es precisamente estudiar el proceso de aprendizaje de Ingeniería Informática en la municipalización.

En tal sentido, para la realización del estudio se seleccionó el mayor conjunto posible de estudiantes que hubiesen cursado el mayor número de asignaturas posible, siendo 100 los estudiantes que cumplen esta condición para un total de 36 asignaturas. Frente a la posibilidad de que un estudiante tuviese varias evaluaciones en una misma asignatura, según el número de veces que cursó ésta, se seleccionó la última evaluación registrada correspondiente a la última vez que cursó la asignatura, coincidiendo en este caso que los 100 estudiantes poseen aprobadas las 36 asignaturas, por lo que el conjunto de resultados oscila entre los valores 3 y 5. Si bien el número de veces que un estudiante ha repetido una asignatura constituye un elemento de gran importancia en el presente estudio, el mismo se tendrá en cuenta como un fenómeno latente de gran influencia en los patrones de comportamiento de los estudiantes, y la forma de reflejarlo en el análisis será tratado en estudios posteriores, siendo el actual un primer acercamiento al tema.

B. Métodos

La transformación de las calificaciones de los estudiantes en conocimiento se efectuó por medio de la aplicación de técnicas de análisis multivariado (específicamente análisis de componentes principales, análisis de *clusters* y análisis factorial común), las cuales se encuentran descritas para variables numéricas en escala de intervalos; mientras que las calificaciones en la Enseñanza Superior Cubana se emiten en una escala ordinal, semicuantitativa, sin intervalo de magnitud definido e igual entre cualquier par de niveles consecutivos de la variable [6].

Como paso previo a la realización de los análisis, fue necesario desarrollar un proceso de transformación de las notas de los estudiantes en las diferentes asignaturas, obteniéndose datos transformados que pueden tomar valores numéricos desde el 1 hasta el 100; con el objetivo de atenuar el impacto que sobre el análisis posee el hecho de que las notas son variables discretas y de otros fenómenos asociados al valor que posee y representa una nota en específico, en el que, por ejemplo, un examen calificado con 2 no significa necesariamente que es la mitad de bueno que uno que fue calificado con 4. En este sentido, los datos transformados van a representar el número de orden que le corresponde al alumno dentro del conjunto evaluado en una asignatura determinada, dado que el 100 representa al alumno que dentro de este total obtuvo la mejor calificación y el 1 representa al que obtuvo la peor. Es recomendable realizar este tipo de transformación siempre que se realicen análisis multivariados con datos ordinales [7].

El análisis de la relación entre las variables (elemento indispensable para la realización de la reducción de las mismas minimizando y/o controlando la pérdida de información) se efectuó a partir del estudio de las matrices de correlación, tanto de Pearson (realizada a las notas de los estudiantes) como de Spearman (realizada a la transformación que se le efectuó a dichas notas) y de la matriz de dispersiones. Se incluyeron en el análisis las correlaciones de Spearman, ya que las variables poseen escala ordinal; sin embargo, también se tuvieron en cuenta las correlaciones de Pearson. Se empleó la matriz de dispersiones con el objetivo de, a partir de su análisis gráfico, detectar posibles relaciones no lineales entre las variables, elemento éste que provocaría que los resultados arrojados, tanto por el análisis de componentes principales, como por el análisis factorial común, fueran no confiables y, por tanto, inservibles. El referido análisis se realizó, tanto para los datos originales, como para los datos transformados, con el objetivo de corroborar los resultados y las decisiones que a partir de éstos fueron tomadas [8].

A partir de las técnicas anteriormente mencionadas, se decidió conservar en el estudio las 36 asignaturas iniciales, a las que se les aplicó el análisis de componentes principales, con el objetivo de reducir el número de dimensiones y por tanto, la cantidad de datos. Así, se obtuvo un conjunto reducido de variables independientes que explican un porcentaje significativo de la varianza total y caracterizan los resultados de cada estudiante.

Para la determinación del número de componentes principales a conservar, se emplearon criterios asociados al análisis de los autovalores. Fueron desechadas las componentes que poseen autovalores inferiores a 1, puesto que no son capaces de explicar al menos una variable. De igual forma, se empleó la representación gráfica de las componentes y sus respectivos autovalores, identificándose la componente a partir de la cual la curva comienza a rectificarse y desechándose aquellas componentes asociadas a la mencionada rectificación de la curva [8]; de este modo se conservó un conjunto mínimo que explica el 62 % de la varianza total, ya que se trata de una aplicación de ciencias sociales.

A las componentes seleccionadas se les investigó la normalidad mediante la dócima de Anderson-Darling, y las últimas componentes, que expresan la influencia de las fuentes aleatorias de variabilidad, se usaron para identificar alumnos con comportamientos atípicos [8].

Se le aplicó la técnica de *clusters* jerárquicos [9] a las 4 componentes principales, con el objetivo de identificar grupos de estudiantes con comportamientos similares. Para ello se emplearon distancias Euclidianas y de Manhattan [9] y se probaron enlaces por Centriodes y de Ward [8] para comparar los resultados y seleccionar la combinación que más se ajusta a los objetivos del trabajo, seleccionando para el análisis la variante con mayor porcentaje de similaridad y menor número de grupos [8]. Los resultados se emplearon para representar a los estudiantes según el grupo al que pertenecen en el gráfico de las componentes 1 y 2, y en el gráfico de las componentes 3 y 4; y analizar dicho agrupamiento con el objetivo de, a partir del ordenamiento de los datos originales de los estudiantes según el grupo al que pertenecen, comprobar las similitudes existentes entre los miembros de un mismo grupo y las diferencias que existen entre los grupos identificados.

ANÁLISIS MULTIVARIADO DEL APRENDIZAJE DE INGENIERÍA INFORMÁTICA EN LA MUNICIPALIZACIÓN

Con el objetivo de identificar los factores latentes que reflejan lo que las variables tienen en común, se le realizó a las componentes principales al análisis factorial común, como técnica de ayuda en la identificación de regularidades en el comportamiento de los estudiantes y las posibles causas que las provocan. Se empleó una rotación varimax para obtener una mejor representación en un solo factor de cada una de las asignaturas y las cargas factoriales asociadas se utilizaron como representación de la influencia que tiene cada componente sobre cada una de las asignaturas empleadas en el estudio [8].

III. RESULTADOS

El análisis de componentes principales con 36 asignaturas permitió definir 4 componentes principales (variables) que condensan el 62 % de la variabilidad total de las evaluaciones obtenidas por los estudiantes en las referidas asignaturas. Las mismas, además, son capaces de explicar por sí solas más de una variable. Para lograr representar el 90 % de la variabilidad total serían necesarias 15 componentes principales, lo que demuestra que esta proporción de la variabilidad total no puede ser condensada en un número pequeño de variables independientes, fenómeno que se explica por el hecho de que las correlaciones entre las variables no son fuertes. A las 4 componentes principales seleccionadas como significativas, se les asignó un significado de acuerdo a las asignaturas que las caracterizan. Esto se efectuó por medio del análisis de la ubicación de las asignaturas o grupos de éstas en los gráficos de las cargas factoriales de dichas componentes, los cuales se muestran en las figuras 1 y 2.

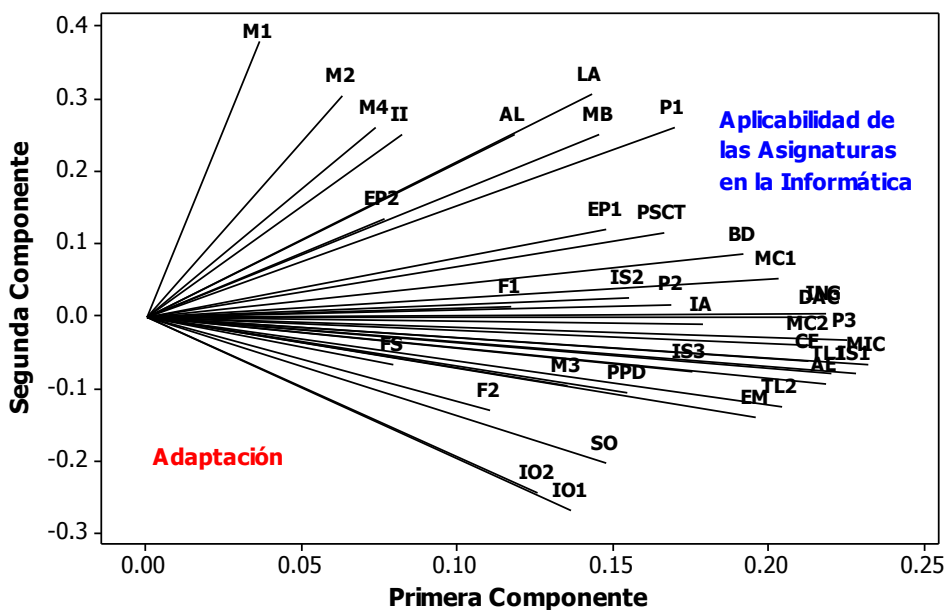


Figura 1. Análisis de las cargas de las asignaturas en las dos primeras componentes principales.

Los diferentes análisis de *clusters* realizados, empleando distancias Euclidianas y de Manhattan y probando enlaces por Centroides y de Ward, arrojaron que la variante de enlace por Centroides con distancia Euclidiana es la que más se acerca al criterio de aceptación, ya que consistió en un corte del dendrograma con 13 grupos, con un 80 % de nivel de similaridad dentro de un grupo y una distancia de 2,56. Ello permitió obtener como resultado 4 grupos de buen aprendizaje (donde se concentran solamente 16 estudiantes), 2 grupos de rendimiento medio (compuestos solamente por 4 estudiantes en total) y 7 grupos con dificultades (donde se agrupan los 80 estudiantes restantes, de ellos 62 se corresponden con el Grupo 3, que a su vez constituye el grupo con mayores dificultades). Cada uno de los citados grupos representa un estado diferente de aprendizaje, cuya interpretación está dada por su ubicación en los ejes factoriales [8], como puede apreciarse en los gráficos de las figuras 3 y 4. Así, por ejemplo, los grupos 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.

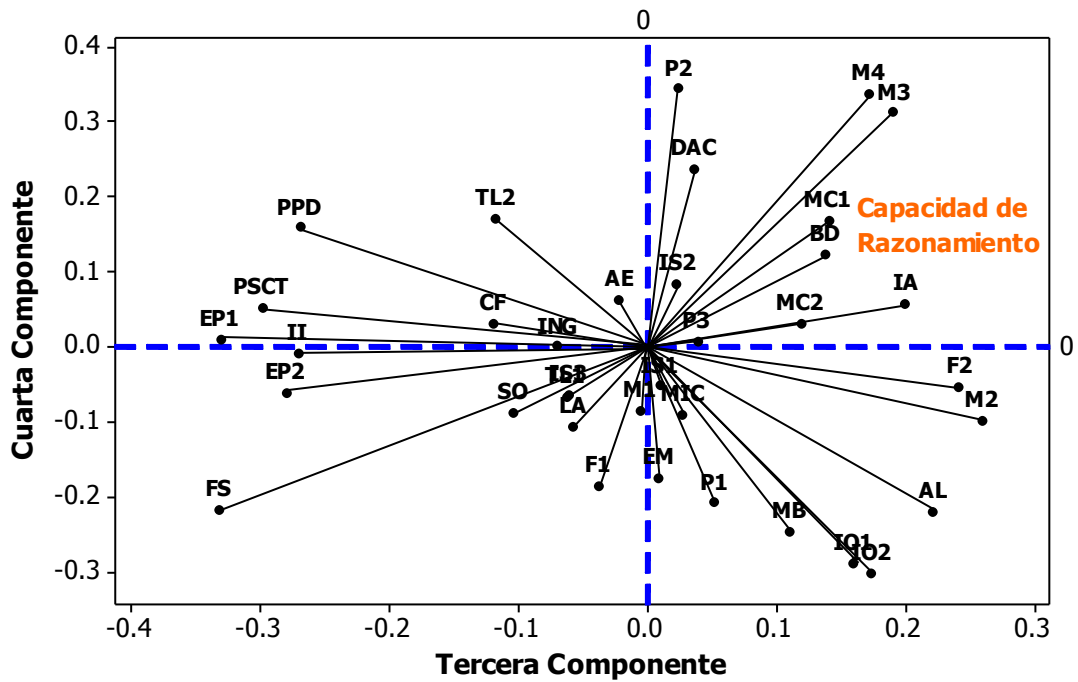


Figura 2. Análisis de las cargas de las asignaturas en la tercera y cuarta componentes principales.

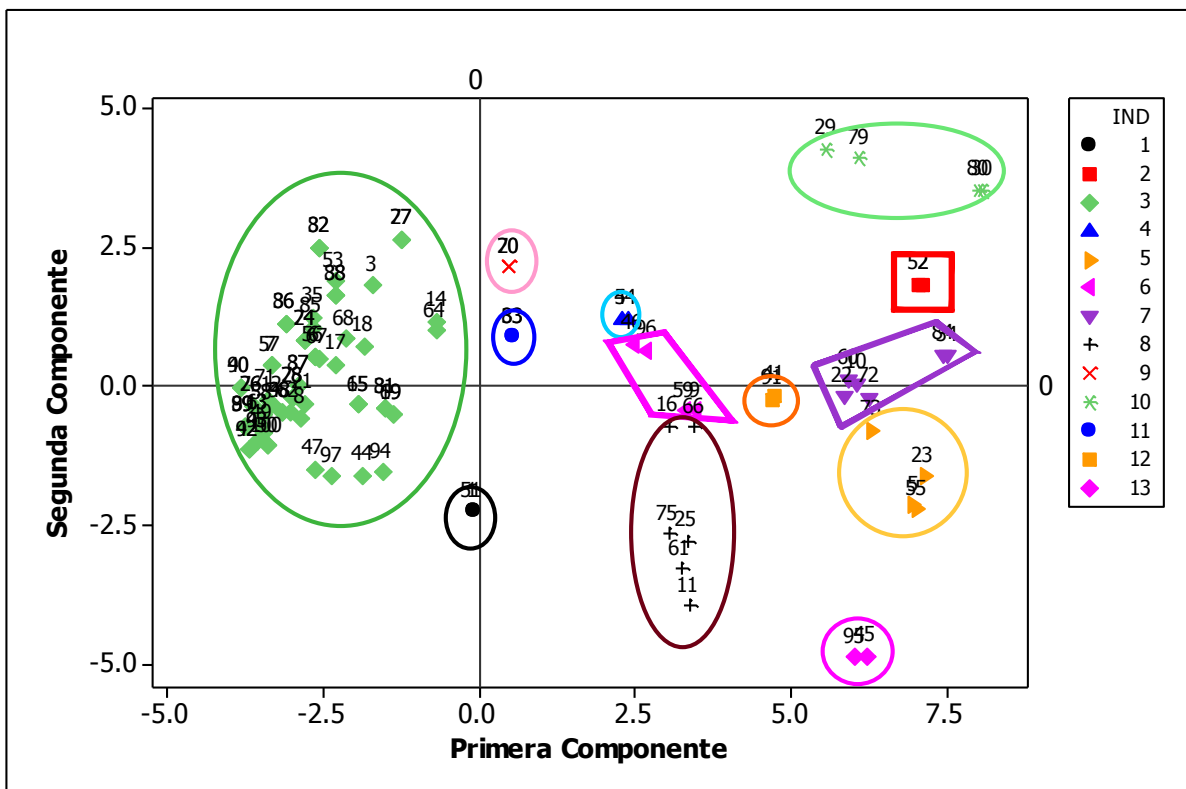


Figura 3. Representación de las componentes 1 y 2 obtenidas en el Análisis de *Clusters* con enlace por Centroides y distancia Euclidiana.

ANÁLISIS MULTIVARIADO DEL APRENDIZAJE DE INGENIERÍA INFORMÁTICA EN LA MUNICIPALIZACIÓN

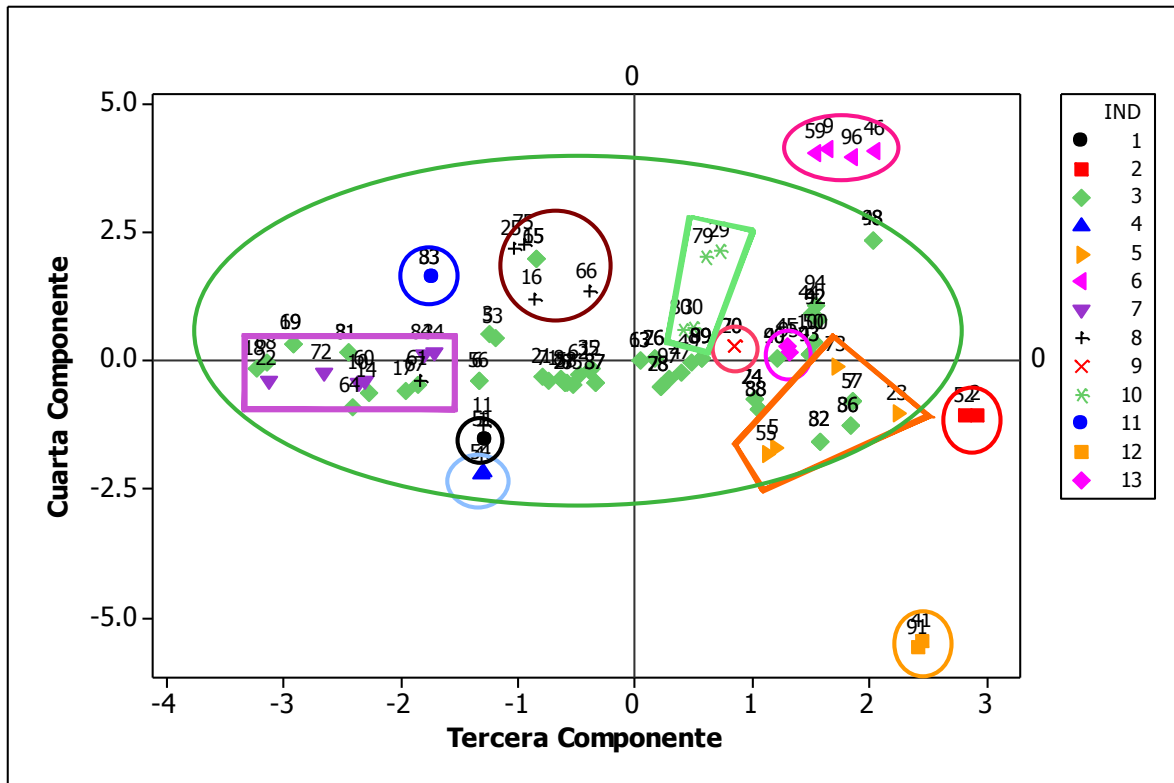


Figura 4. Representación de las componentes 3 y 4 obtenidas en el Análisis de Clusters con enlace por Centroides y distancia Euclidiana.

En la realización del análisis factorial común se empleó una rotación Varimax, obteniéndose 5 factores latentes que determinan el comportamiento de las 36 asignaturas incluidas en el estudio sin necesidad de eliminar ninguna y cuyas representaciones se pueden apreciar en los gráficos de las figuras 5, 6 y 7.

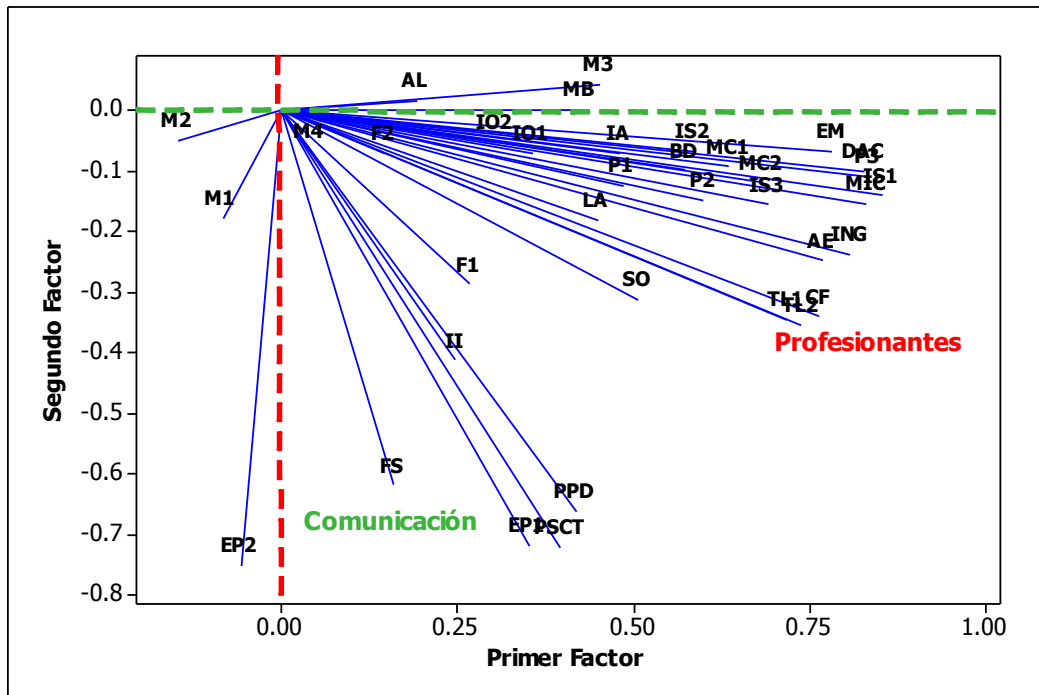


Figura 5. Representación del primer y segundo factor obtenidos en el Análisis Factorial Común empleando Rotación Varimax.

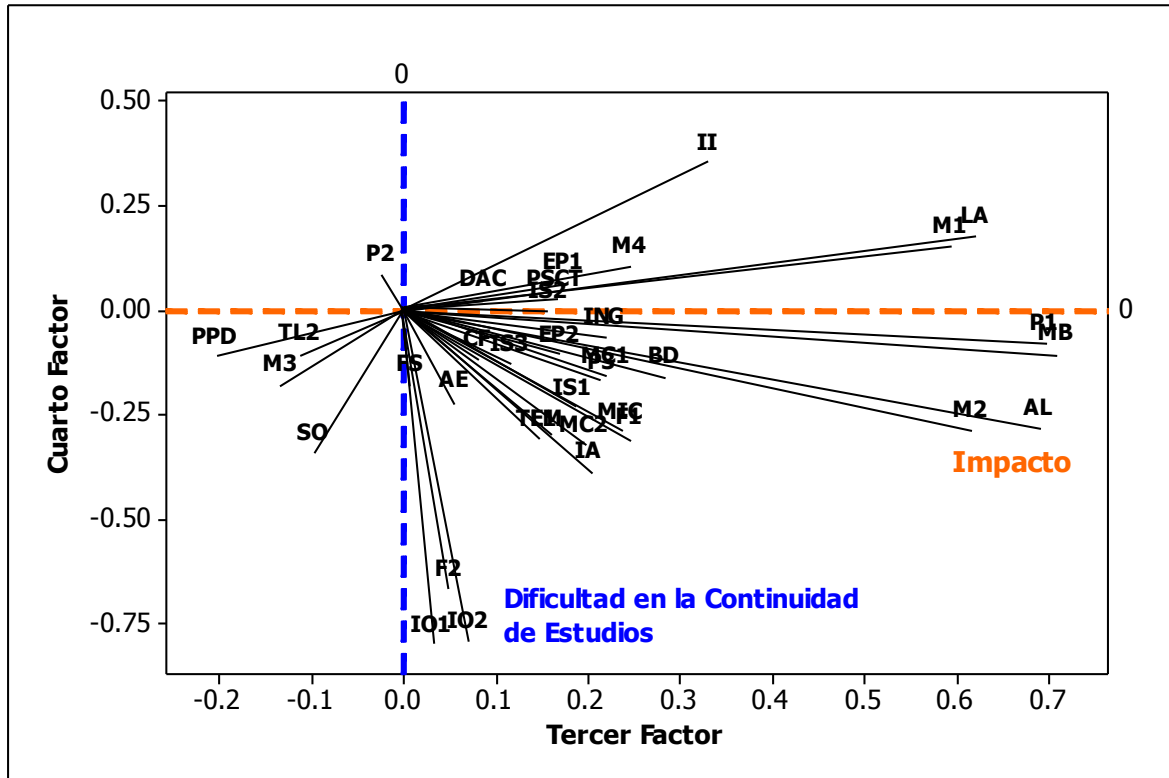


Figura 6. Representación del tercer y cuarto factor obtenidos en el Análisis Factorial Común empleando Rotación Varimax.

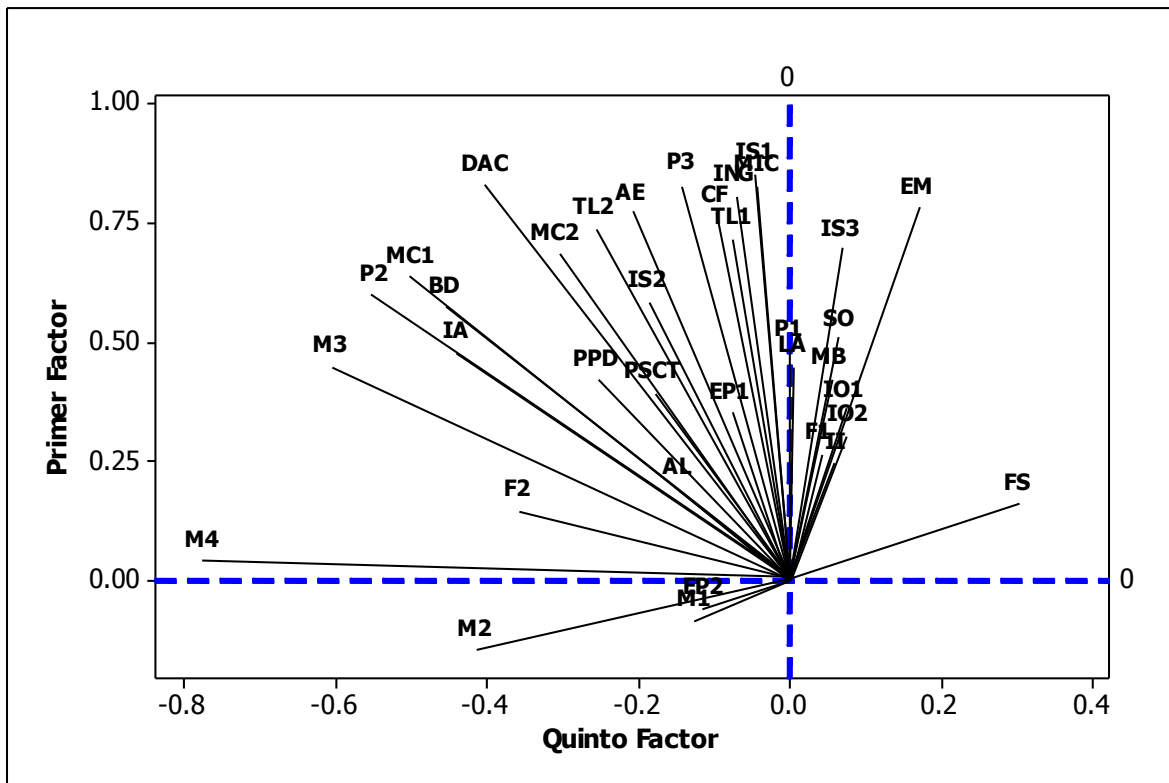


Figura 7. Representación del primer y quinto factor obtenidos en el Análisis Factorial Común empleando Rotación Varimax.

IV. DISCUSIÓN

Para el logro de la mejor calidad posible en el proceso de enseñanza-aprendizaje, resulta vital tener en cuenta (como elemento fundamental en la gestión de éste) la variabilidad que representan las individualidades de los estudiantes. En tal sentido, las acciones desarrolladas por los profesores sobre los alumnos juegan un papel fundamental, independientemente de que se conoce que la efectividad de éstas depende de las características propias de los estudiantes (que son muy variadas), de sus motivaciones internas, de sus habilidades para aprender; en fin, de las posibilidades de desarrollo alcanzado por el estudiante en el momento del aprendizaje. Es por tanto esencial, que quienes guían al alumno mediante el citado proceso, posean una evaluación razonable de sus posibilidades y le orienten acerca de cómo aprovecharlas y estimularlas en aras de obtener el mejor rendimiento académico posible.

El profesor debe contar, entonces, con una evaluación sintética del rendimiento académico del estudiante, la cual debe generarse a partir de una estrategia de análisis de datos para el descubrimiento del conocimiento asociado al proceso de enseñanza [10]. Esta evaluación debe poseer, además, la menor pérdida posible de información cualitativa, ya que, por el contrario, el profesor tendría que manejar un cúmulo elevado de información que limitaría la realización de su trabajo. En tal sentido, las técnicas de análisis multivariado juegan un papel fundamental, ya que uno de sus campos de estudio es precisamente la reducción del número de variables a partir de un grupo mayor de éstas, de manera tal que se obtenga la menor pérdida posible de información [11].

Teniendo en cuenta los elementos anteriormente planteados, se propone la obtención de "clusters de estudiantes" como vía para evaluar el rendimiento académico de éstos (tomando como referencia sus calificaciones en las diferentes asignaturas), de forma tal que se logren identificar las características individuales de cada estudiante mediante la caracterización del grupo 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 [12; 13], ya que éstas igualan numéricamente los resultados académicos de estudiantes que poseen evaluaciones muy diferentes, si se tiene en cuenta el tipo de materia, perdiéndose gran cantidad de información cualitativa. Esto limita la posibilidad de encontrar factores que expliquen dicho comportamiento en los resultados de los estudiantes. Por otra parte, en muchas universidades del mundo, el alumno matricula por curso un número de materias de su elección, siguiendo determinadas reglas de precedencia, por lo que unos pueden graduarse en menor período que otros, lo cual crea la necesidad de tener en cuenta, no sólo las calificaciones obtenidas, sino también la cantidad de materias por curso y algún criterio de celeridad para la síntesis de la calidad de los resultados en una o pocas variables [14].

Se propone, además, realizar el agrupamiento de los estudiantes a partir de la obtención de un mínimo número de variables independientes (componentes) que sinteticen los resultados de los estudiantes en las diferentes asignaturas, lo cual se hace a partir de un análisis de componentes principales y de un análisis factorial común. La citada reducción de variables se realiza por lo impracticable que resultaría emplear como base los resultados del total de asignaturas cursadas por los estudiantes, ya que al ser éstas una cantidad elevada, dificultaría asignarle un significado a los *clusters* obtenidos.

A. Análisis de Componentes Principales

El análisis del gráfico de la figura 1 muestra como en su eje x (correspondiente a la primera componente) todas las asignaturas cargan en el mismo sentido, lo que indica que todas son capaces de discriminar a los estudiantes aunque unas lo son más que otras. Las asignaturas que se ubican más a la derecha son las que poseen mayor capacidad para diferenciar a los estudiantes, coincidiendo en este caso con las asignaturas más relacionadas con el perfil del Ingeniero Informático. Pudiera inferirse entonces que la primera componente se encuentra asociada a la aplicabilidad de las materias en el ejercicio de la profesión. Por otra parte, las materias asociadas a los primeros años de la carrera se encuentran ubicadas en la parte superior (como los Cálculos, Álgebra, etc.) y en la parte inferior las asociadas a la culminación de estudios (como Investigación de Operaciones). Este fenómeno pudiera indicar que sobre la segunda componente influye la adaptación de los estudiantes al tipo de enseñanza y a los nuevos enfoques y retos que les plantean las asignaturas a medida que van siendo cursadas en los diferentes años. Pudiera interpretarse, en otras palabras, como un proceso de formación de hábitos de conducta o

pensamiento, que, en dependencia de su estadío, resultan favorables para algunas materias y para otras no.

Los estudiantes de los Centros Universitarios Municipales poseen diverso origen y en su mayoría carecen de la continuidad de estudios, por lo que no es de extrañar que nuevas asignaturas con nuevos enfoques analíticos impacten en dichos estudiantes con mayor fuerza que otras que ya poseen precedentes, siendo las primeras más capaces de discriminarlos. Este fenómeno posee mayor incidencia en los primeros años de la carrera, cuando aún están en formación muchas de las habilidades que se incorporan y la disciplina y constancia en el estudio. Nótese además cómo las asignaturas asociadas al perfil profesional de la informática, se encuentran cerca del 0 del eje y, ubicándose una parte importante de ellas por debajo de dicha línea, lo cual indica que el impacto que estas asignaturas poseen sobre el estudiante es menor, ya que muchas de ellas poseen asignaturas precedentes en las que ya sufrieron procesos de adaptación. En este sentido, se observa cómo en Programación 1, los estudiantes sufren un impacto casi tan fuerte como en los Cálculos y Álgebra; poseyendo, por medio de este enfoque, una capacidad mucho mayor de discriminarlos que Investigación de Operaciones, que aunque ofrece nuevos enfoques, se ubica en los momentos terminales de la carrera, encontrándose los estudiantes mucho más preparados para asumirla.

En el caso de las componentes principales tercera y cuarta (cuya representación corresponde a la figura 2), condensan solamente el 13,1 % de la variabilidad total. Al analizar la tercera componente el gráfico muestra agrupadas a su izquierda a las asignaturas: Economía Política, Preparación para la Defensa, Filosofía y Sociedad, Introducción a la Ingeniería y Problemas sociales de la Ciencia y la Tecnología; encontrándose el resto de las asignaturas ubicadas gradualmente según crece el eje x. Este fenómeno ofrece indicios de que la tercera componente se encuentra influenciada por las habilidades que las asignaturas requieren en cuanto a capacidad de razonamiento, siendo Matemática 2 la que más capacidad posee en este sentido de discriminar a los estudiantes. En cuanto a la cuarta componente, se aprecia un contraste evidente entre los resultados de asignaturas Cálculo 3 y 4 (M3 y M4) e Investigación de Operaciones (IO1 e IO2); en este caso no ha sido evidente caracterizar cada uno de los subgrupos de asignaturas que se separan. No obstante, esto no disminuye el valor de los resultados obtenidos, pues lo esencial es conocer que en el período estudiado hubo subconjuntos de alumnos con resultados opuestos en estos grupos de materias.

Las cuatro componentes principales determinadas con sus respectivas interpretaciones simplifican en gran medida la valoración de los estudiantes a partir de sus resultados docentes y permiten identificar factores latentes en el proceso de enseñanza-aprendizaje. Por ejemplo, se puede conocer cómo se comporta en un estudiante la formación de habilidades asociadas a la capacidad de razonamiento, o si el alumno se está adaptando de forma correcta a los nuevos enfoques y retos que les plantean las asignaturas, etc.; lo cual le permite al profesor ejecutar las acciones que el estudiante verdaderamente requiere para alcanzar el mejor rendimiento académico posible. Otra de las bondades de la técnica lo constituye la identificación de estudiantes atípicos, es decir, aquellos estudiantes en los que no son aplicables la mayoría de las reglas y asociaciones de comportamiento de resultados que pueden imputarse a determinados grupos de estudiantes frente a las asignaturas. Esto le permite al profesor investigar las causas del comportamiento de dichos alumnos y generar estrategias de ayuda.

B. Análisis Factorial Común

El análisis factorial común, al igual que el análisis de componentes principales, permite sintetizar en pocas variables la información contenida en un número mayor de éstas con la menor pérdida de información posible, pudiéndose interpretar de forma similar sus resultados. En este caso, dicho análisis se empleó con el objetivo de complementar los resultados obtenidos a través del análisis de componentes principales y fortalecer la identificación de factores latentes que reflejan lo que las variables tienen en común, así como la identificación de regularidades en el comportamiento de los estudiantes y las posibles causas que las provocan. Esto ofrece mayor certeza en los resultados y garantiza que las acciones de los profesores sobre los alumnos se correspondan con las verdaderas necesidades de estos últimos y, por tanto, se potencie el incremento del rendimiento académico. Sin embargo, no resulta absolutamente necesario complementar entre sí los resultados de los mencionados análisis, ya que cada uno de ellos es capaz, por sí mismo, de cumplir eficientemente el objetivo que con su realización se pretende alcanzar.

ANÁLISIS MULTIVARIADO DEL APRENDIZAJE DE INGENIERÍA INFORMÁTICA EN LA MUNICIPALIZACIÓN

Como se puede observar en la figura 5, en el primer factor del Análisis Factorial Común se agruparon 18 asignaturas, resultando ser las más vinculadas al perfil principal del Ingeniero Informático, asignaturas "Profesionantes", incluyendo otras de apoyo como Inglés, Contabilidad y Finanzas, Administración de Empresas y Metodología de la Investigación Científica. En el segundo factor se agruparon 6 asignaturas de corte comunicativo, relacionadas con las Ciencias Sociales, las cuales generalmente al estudiante de esta especialidad le resultan menos atractivas, incluyéndose aquí la Introducción a la Ingeniería y la Preparación para la Defensa.

El análisis del gráfico de la figura 6 muestra cómo, en el tercer factor, se agruparon las 6 asignaturas más difíciles del primer año de la carrera, que podrían llamarse asignaturas de "Impacto". Esto se debe a que, fundamentalmente los estudiantes de este tipo de enseñanza, son trabajadores, que cuando se insertan en los Centros Universitarios Municipales, les resulta complicado el modelo semipresencial de estudios, sobre todo en estas asignaturas que demandan muchas horas de estudio y ejercitación; y que a pesar de ser la base de la Ingeniería, para ellos no constituye incentivo y muchos se cuestionan incluso su impartición. Por otra parte en el cuarto factor se agruparon 4 asignaturas que para los estudiantes resultan ser las más complejas de la continuidad de estudios, pues los hacen llevarlas como arrastre en años posteriores. Además, llegan hasta frenarles la presentación del trabajo de culminación de estudios o tesis final, pues aunque las materias de Física son de tercer año, al no ser precedentes para otras asignaturas de los años posteriores, las pueden seguir matriculando en los años siguientes; no así las materias de Investigación de Operaciones, que se imparten en quinto y sexto año de la carrera.

Finalmente, el gráfico de la figura 7 muestra cómo en el quinto factor, se agruparon 2 asignaturas (Matemática 3 y 4), que aunque tienen gran semejanza con el tercer factor, quedan independientes, pues ya se dan en el segundo año de la carrera, donde el alumno ya se está adaptando al modelo semipresencial de estudios, resultándole menos complicadas.

C. Análisis de Clusters

El análisis de *cluster* realizado muestra cómo en los grupos con dificultades se agrupa el 80 % de los alumnos estudiados, lo cual resulta preocupante. El Grupo 3 constituye el de mayor dificultad, agrupándose en éste el 62 % de los estudiantes. Como se puede observar en la figura 3, el Grupo 3 se ubica a la izquierda de la primera componente, lo cual indica que los estudiantes que lo componen poseen bajas calificaciones en todas las asignaturas. De igual forma puede observarse (tanto en la figura 3 como en la figura 4) cómo estos estudiantes poseen dificultades en la formación de hábitos de conducta, estilos de pensamiento y habilidades de razonamiento que les permitan asimilar los nuevos enfoques y retos que les plantean las asignaturas a medida que van siendo cursadas en los diferentes años. Se considera que en alguna medida este fenómeno posee relación con las facilidades en cuanto a repitencia y arrastre que ofrecía el modelo de enseñanza empleado en los Centros Universitarios Municipales en el período estudiado, ya que no se limitaba el número de veces que se podía repetir un año o cursar una misma asignatura. Estas facilidades, unidas al hecho de que los alumnos son también trabajadores, provocan que sea menor la presión y necesidad de incorporación del conocimiento, ya que en caso de suspender una asignatura ésta puede cursarse nuevamente.

Una vez culminado el análisis anterior, correspondería la elaboración de estrategias para potenciar el aumento del rendimiento docente de los estudiantes, las cuales deberían enfocarse en la solución de los elementos identificados como causas de los problemas que se poseen. Resulta importante señalar, cómo las modificaciones desarrolladas a partir del 2009 por el Ministerio de Educación Superior en los reglamentos que regulan la enseñanza en los Centros Universitarios Municipales, se encuentran alineadas a los resultados obtenidos en el presente estudio. Dichas modificaciones se centran en limitar las posibilidades de repitencia y arrastre de asignaturas por parte de los estudiantes, así como en la exigencia de la asistencia a clases, con el objetivo de potenciar el rendimiento docente; lo cual, en alguna medida, valida los resultados obtenidos en el presente estudio y la utilidad de las técnicas de análisis multivariado para la mejora del proceso de enseñanza-aprendizaje.

En el caso de los grupos de rendimiento medio, éstos se encuentran representados por los Grupos 12 y 13. Estos grupos se diferencian de aquellos que tienen dificultades, ya que poseen resultados significativamente superiores en las asignaturas básicas (Álgebra, Física 1 y 2, Matemática 1 y 2, etc.) y las terminales como Investigación de Operaciones. El Grupo 13 se diferencia del 12 en que los resultados de los estudiantes en las asignaturas asociadas al perfil de la carrera y a las Ciencias Sociales (Filosofía y Sociedad, Economía Política, Problemas Sociales de Ciencia y la

Tecnología, Preparación para la Defensa), son superiores, no ocurriendo así en el resto de las asignaturas.

Los grupos clasificados de buen aprendizaje se diferencian de los de rendimiento medio en que sus estudiantes, en sentido general, poseen mejores calificaciones en las asignaturas. Estos grupos se diferencian entre sí en que unos poseen mejores resultados que otros en las asignaturas de Ciencias Sociales, como es el caso del Grupo 7 respecto al 5. El Grupo 10 posee mejores resultados que el 5 y el 6 en la asignatura Investigación de Operaciones, además de poseer a los 2 estudiantes de mejores resultados docentes. En el caso del Grupo 2, éste se diferencia del Grupo 10 en que sus resultados en las Ciencias Sociales son ligeramente inferiores. La caracterización de los grupos determinados y, por tanto, de los estudiantes que los componen, ofrece excelentes oportunidades para orientar el trabajo de los profesores con los estudiantes en la búsqueda de elevar la calidad del proceso docente.

V. CONCLUSIONES

1. A partir de los resultados de los estudiantes, se lograron identificar las relaciones que existen entre las diferentes asignaturas y, por tanto, los factores latentes que influyen sobre el proceso de aprendizaje.
2. Se logró sintetizar el desempeño de los estudiantes en un solo indicador, pudiendo ser éstos identificados a partir de las características de su *cluster* de pertenencia, las cuales se asocian a las materias en las que poseen mayores éxitos o insuficiencias.
3. El análisis de componentes principales y el análisis factorial común pueden ser empleados como complementos entre sí, con el objetivo de fortalecer sus resultados; sin embargo, esto no resulta absolutamente necesario.
4. La caracterización presentada del aprendizaje de los estudiantes, se centra en la evaluación de sus conocimientos y habilidades según las calificaciones, no teniéndose en cuenta en dichos resultados la valoración del desarrollo del pensamiento y la formación de valores lograda por el alumno; por tanto, para lograr una valoración adecuada de la calidad del proceso de aprendizaje, deben integrarse todos los aspectos anteriormente comentados.
5. Las técnicas de análisis multivariado aplicadas permitieron identificar elementos del modelo de enseñanza empleado que limitan la calidad del proceso docente, como lo es la repitencia y arrastre de asignaturas de forma indefinida, y por tanto, ofrecen la posibilidad de corregir y mejorar el proceso de forma continua.
6. A partir del conocimiento alcanzado acerca de las características y elementos comunes entre las asignaturas y entre los distintos tipos de estudiantes, de acuerdo a los resultados obtenidos por éstos en el período analizado, se puede realizar el seguimiento de estos alumnos de forma mucho más eficiente y determinar el impacto y características no visibles del modelo de estudios semipresencial aplicado en la municipalización; y por tanto, las divergencias entre sus objetivos y el verdadero cumplimiento de los mismos, para apoyar la toma de decisiones en la corrección de dichas desviaciones y formar un ingeniero cada vez mejor. 🏛️

VI. REFERENCIAS

1. WINOGRAD, M., *Marco Conceptual para el Desarrollo y Uso de Indicadores Ambientales y de Sustentabilidad para la Toma de Decisiones en Latinoamérica y el Caribe* [en línea], Proyecto CIAT/UNEP (International Center for Tropical Agriculture / United Nations for Tropical Environment Programme), 1995 [consulta: Disponible en: <<http://webapp.ciat.cgiar.org/indicators/indicadores/publications.htm>>]
2. RODRÍGUEZ, A.; HEREDIA, J., «Rediseño de procesos de gestión de la enseñanza basado en tecnologías informativas», en *Novena Semana Tecnológica: Las TIC presente y futuro* La Habana, Ministerio de la Informática y las Comunicaciones (MIC), 2009, [consulta: 2010-12-04]. Disponible en: <<http://www.dtic.co.cu/novenasemanatecnologica>>
3. FDA, *Guidance for Industry: PAT - A Framework for Innovative Pharmaceutical Development, Manufacturing and Quality Assurance* [en línea], U.S Department of Health Services, Food and Drug Administration, 2004 [consulta: 2010-12-08]. Disponible en: <<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm070305.pdf>>

ANÁLISIS MULTIVARIADO DEL APRENDIZAJE DE INGENIERÍA INFORMÁTICA EN LA MUNICIPALIZACIÓN

4. DAVENPORT, T.; SHORT, J., «The New Industrial Engineering: Information Technology and Business Process Redesign» *Sloan Management Review*, 1990, vol. 31, no.4, pp. 11-27, ISSN 1532-9194.
5. LUAN, J., «Data Mining an Knowledge Management in Higher Education», en *Annual Forum for the Association Institutional Research* Toronto (Canada), 2002, [consulta: 2010-12-08]. Disponible en: <<http://www.eric.ed.gov>>
6. RODRÍGUEZ, A. G.; HAEDO, Y., «Análisis Multivariado del Proceso de Aprendizaje en el Primer Año de Ingeniería», en *14 Convención Científica de Ingeniería y Arquitectura. V Simposio de Ingeniería Industrial y Afines. V Taller de Enseñanza de la Ingeniería Industrial* (1 al 5 de diciembre), Palacio de Convenciones de La Habana, Instituto Superior Politécnico José Antonio Echeverría, Cujae, 2008, ISBN 978-959-261-281-5.
7. MANISERA, M.; VAN DER KOOJI, A.; DUSSELDORP, E., «Identifying the Component Structure of Satisfaction Scales by Nonlinear Principal Components Analysis» *Quality Technology & Quantitative Management*, 2010, vol. 7, no. 2, pp. 97-115, ISSN 1684-3703.
8. HAIR, J. F.; ANDERSON, R. E.; TATHAM, R. L.; BLACK, W. C., *Análisis Multivariante*, 5ta. ed., Madrid, Prentice Hall Iberia, 1999, ISBN 8483220350, pp. 92-94.
9. JOHNSON, D. E., *Métodos Multivariados Aplicados al Análisis de Datos*, México, D. F., International Thomson Editores, 2000, ISBN 9789687529905.
10. DELAVARI, N.; REZA, M., «Data Mining Application in Higher Learning Institutions» *Informatics in Education*, 2008, vol. 7, no. 1, pp. 31- 54, ISSN 1648-5831.
11. JOHNSON, R. A.; WICHERN, D. W., *Applied Multivariate Statistical Analysis*, 4ta. ed., New Jersey, Prentice Hall, 1998, ISBN 10 013834194X, pp. 541-544.
12. CHOW, T., «Predicting End of 1st Year Performance», en *Annual Forum for the Association for Institutional Research* Boston, 2004, [consulta: 2010-11-15]. Disponible en: <<http://www.air2004.org>>
13. MORAL, J., «Predicción del rendimiento académico universitario» *Perfiles Educativos*, 2006, vol. 28, no. 113, pp. 38-63, ISSN 0185-2698.
14. GARNICA, E., «El Rendimiento Estudiantil: Una Metodología para su Medición», *Economía* [en línea], 1997, no. 13, pp. 7-26 [consulta: 2010-12-09], ISSN 1315-2467. Disponible en: <<http://www.saber.ula.ve/handle/123456789/19178>>