

ESTUDIO DE LAS VARIABLES DEL MODELO DE INDICADORES DE GESTION DE LAS UNIVERSIDADES PÚBLICAS COLOMBIANAS

RESUMEN

El estudio de las variables del modelo de indicadores de gestión del sistema de universidades públicas (SUE) desde el punto de vista institucional, se convierte en un estudio de carácter prioritario, ya que actualmente el ministerio de educación asigna los recursos a distribuirse entre las universidades públicas con base en estos indicadores. Desafortunadamente muchas universidades colombianas no cuentan con estudios de dichos indicadores, reduciendo la posibilidad mejorar los aspectos necesarios para tener acceso a mayores recursos que permitan su funcionamiento óptimo.

En este artículo se analizan las variables que inciden dentro del modelo de indicadores de gestión y se propone una metodología general basada en técnicas de inteligencia artificial que ayudará a elaborar un modelo matemático para determinar los ajustes necesarios que debe hacer cada institución con el fin de mejorar sus índices de gestión.

JOHN JAIRO SANTA

Ingeniero Electricista, M. Sc.
Profesor Auxiliar
Universidad Tecnológica de Pereira
jjsanta@utp.edu.co

WILLIAM ARDILA URUEÑA

Licenciado en Física
MSc. Física
Profesor Titular
Universidad Tecnológica de Pereira
williamar@utp.edu.co

PALABRAS CLAVES: Modelo matemático, sistema de gestión de universidades públicas, técnicas de inteligencia artificial, variables modelo

ABSTRACT

The study of variables of the public universities (SUE) management indexes models from the institutional point of view has become in an priority study due the national education department assigns economical recourses to the colombian universities based on that indexes. Unfortunately several universities have not studies about those indexes, reducing the possibility to improve necessary aspects to have access to major resources that allow its optimal operation.

In this paper the most important variables who affect the management indexes model are analyzed and a general methodology based on artificial intelligence techniques are proposed in order to design a mathematical model to determine the necessary adjustments that every university most do in order to improve the management indexes

KEYWORDS: Mathematical model, public universities management indexes, artificial intelligence techniques, model variables

1. INTRODUCCIÓN

De acuerdo con la ley 30 de 1992, las fuentes de recursos de las universidades estatales son tres: los provenientes de la nación, los aportados por entidades territoriales y los propios de cada institución, los cuales se obtienen a través de venta de servicios, matrículas, derechos académicos, programas de extensión, consultorías o convenios de investigación, entre otros.

En cuanto a los recursos provenientes de la nación, estos se aseguran y regulan mediante dos normatividades: los artículos 86 y 87 de la Ley 30. El artículo 86 garantiza a las universidades del estado recursos fijos que provienen de la nación y de las entidades territoriales, los cuales deben ser destinados a funcionamiento e inversión; el artículo 87 por su parte fija el incremento anual de los recursos que aporta la nación por artículo 86 de acuerdo con el crecimiento de la economía y de conformidad con

los objetivos previstos para el sistema de universidades estatales.

Los aportes de la nación provenientes del artículo 86, que constituyen en muchos casos en la principal fuente de financiamiento de las universidades, son distribuidos mediante una ecuación que garantiza su incremento en pesos constantes, tomando como base los presupuestos de 1993. A pesar de lo anterior, dicha ecuación no tiene en cuenta criterios de gestión de las universidades en términos de la eficiencia en el uso de los recursos ó del cumplimiento de las metas institucionales y de la política sectorial.

El gobierno nacional en el anterior plan de desarrollo (Ley 812 de 2003) buscó una modificación al esquema de distribución de estos recursos a partir de la aplicación de un modelo de indicadores de gestión para la asignación de los mismos, el cual fue desarrollado por el sistema de universidades estatales (SUE) y permite fijar criterios

generales para identificar el nivel de eficiencia de las instituciones de educación superior en el manejo y productividad de sus recursos.

A pesar de lo anterior, el artículo 84 de la Ley 812 de 2003 fue declarado inexecutable y por tanto la distribución de los recursos de funcionamiento e inversión de las universidades estatales referidos al artículo 86 de la Ley 30 sigue teniendo su tendencia inercial.

No obstante, la distribución de los recursos correspondientes al artículo 87 de la Ley 30 de 1992, continúa aplicando el modelo de indicadores de gestión diseñado por el sistema de universidades estatales, el cual se ha ido perfeccionando con la participación activa de todas las instituciones que lo componen.

Se trata de un instrumento que combina la estadística y la matemática en un Modelo técnico que provee unos indicadores claros, los cuales permiten a la sociedad enterarse de cómo las universidades estatales producen resultados concretos.

Estos índices están distribuidos en cuatro grandes áreas: Formación, Producción Académica, Bienestar Universitario y Extensión, las cuales son entendidas como las principales áreas misionales de cualquier Institución de Educación Superior (IES). Éstas son evaluadas a partir de la capacidad inicial que posee cada institución, entendida como la combinación de recursos ó insumos con que cuenta la IES para la obtención de sus resultados misionales.

2. INDICES DE CALIDAD

2.1. Índice ICAD

El Índice de Capacidad (ICAD), mide los recursos financieros, humanos, físicos y tecnológicos con los que cuentan las instituciones, y es confrontado con los resultados que obtiene la institución en las cuatro áreas mencionadas, que reflejan la misión de la universidad. La relación entre el índice de capacidad y cada uno de los índices de resultados constituye el marco de referencia para establecer el desempeño de cada universidad en el contexto del Sistema Universitario Estatal. Se trata de un Modelo dinámico, pues contempla un rango estático –el año anterior de la gestión–, y un rango dinámico que incluye vigencias anteriores.

El Modelo de Indicadores de Gestión tiene la virtud de ser acogido por consenso por todas las instituciones que conforman el Sistema de Universidades Estatales, SUE. Este consenso se basa en la transparencia y objetividad del Modelo, en tanto permite que cada universidad se mida “consigo misma”, es decir, de acuerdo con sus propias características y capacidades, comparando aquellos recursos con los que cuenta con los resultados que produce.

Cada institución presenta sus resultados de manera independiente y estos, al ser cotejados con los de otras universidades, arrojan una media que establece una línea de regresión, permitiendo comparar y clasificar los grados de eficiencia de todas las universidades.

Se recoge la información y una comisión trabaja en su depuración y estandarización para entregarla a un grupo de estadísticos que hace el análisis de los resultados. Una vez el modelo es propuesto, éste es revisado y aprobado por el ministerio de educación, para luego hacer la propuesta de distribución de recursos con base en los resultados de gestión producto del modelo utilizado.

Este proceso es observado permanentemente por dos rectores de universidades comisionados por el SUE, de modo que hay un acompañamiento y una comunicación permanente entre el Ministerio y las universidades durante todo el proceso. El SUE también establece una comisión técnica que al final del proceso se reúne con el grupo de estadísticos y revisan los resultados y sus análisis.

La propuesta de distribución de recursos del artículo 87 de la Ley 30 es presentada por parte del Ministerio de Educación Nacional ante el Consejo Nacional de Educación Superior (CESU), máximo órgano asesor y consultivo en materia de Educación Superior. En este punto, el Departamento Nacional de Planeación también tiene acceso a los resultados arrojados por el Modelo, y aunque no tiene una participación activa en el análisis o la ponderación de estos resultados, sus observaciones y sugerencias son tenidas en cuenta para optimizar la asignación de los recursos.

El monto total se distribuye en cuatro bolsas, una para cada grupo de indicadores. Así, por ejemplo, a una universidad puede corresponderle un monto más alto por su gestión en los indicadores de Bienestar, pero uno no tan nutrido en los indicadores de Formación. Las cuatro bolsas se unifican y la suma es el aporte definitivo que le corresponde a cada universidad. Una vez el CESU aprueba esta distribución, el Ministerio de Educación elabora la resolución ministerial a partir de la cual se distribuyen estos recursos

3. TECNICAS UTILIZADAS

3.1. Búsqueda de la proyección

El objetivo ahora es utilizar las variables descritas anteriormente para que sirvan como datos de entrenamiento para uno de los métodos basados en el conocimiento descritos en el numeral B. Sin embargo, como se ha aclarado con anterioridad, el número de variables aquí mostrado es grande (pues algunos describen el comportamiento de cada una de las tres fases) y esto puede ocasionar que el entrenamiento requiera de un esfuerzo considerable. Es por esta razón que se hace importante la reducción de dimensionalidad

de las variables, claro está, sin perder la información que cada una de ellas trae. La Búsqueda de la Proyección (del inglés PROJECTION PURSUIT) [15], el análisis de la matriz de correlación, y el análisis de componentes principales [16], son metodologías que permiten la reducción de la dimensionalidad de un experimento para su mejor análisis.

La PP se realiza a través de la proyección de las variables en diferentes hiperplanos para encontrar el más interesante según el índice de proyección Chi-cuadrado [16]. Este procedimiento se realiza a través de dos etapas:

Búsqueda de la No-normalidad de los datos: El plano es dividido en 48 regiones distribuidas en anillos (figura 2), cada una con un ancho angular de 45° y ancho radial $\sqrt{2 \log 6/5}$, el cual garantiza que cada región tenga la misma probabilidad (1/48) para la distribución normal bivalente. El índice de proyección está dado por

$$PI_{X^2}(\alpha, \beta) = \frac{1}{9} \sum_{j=1}^8 \sum_{k=1}^{48} \frac{1}{C_K} \left[\frac{1}{n} \sum_{i=1}^n I_{B_k} (Z_i^{\alpha(\eta_j)}, Z_i^{\beta(\eta_j)}) - C_K \right]^2$$

C_k : Probabilidad evaluada sobre una región k usando distribución normal bivariada.

n : Número de datos.

I_{B_k} : Es la función indicadora de la región Bk , donde esta es cada las particiones del plano.

Z_i^a : Son las observaciones proyectadas de cada uno de los datos sobre los vectores α y β , donde estos dos últimos son dos vectores ortonormales que son la base del plano de proyección.

n_j : Es el ángulo por el cual el dato es rotado en el plano antes de ser asignado a una de las regiones Bk .

Una de las ventajas del uso del índice Chi-cuadrado es que no se ve afectado en gran forma por outliers.

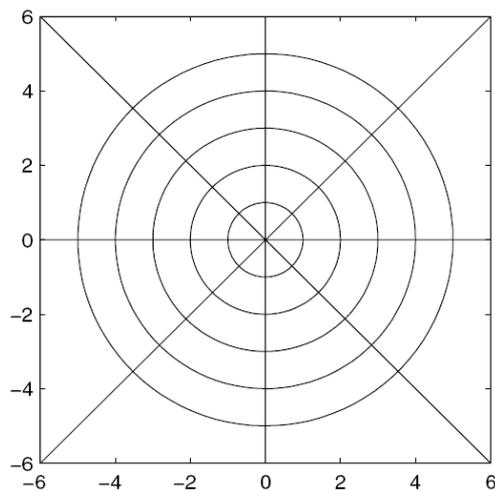


Figura. 2. División del plano donde serán proyectados los puntos del experimento.

El objetivo del PP es encontrar la proyección que da como resultado el mayor de los índices Chi-Cuadrado.

El algoritmo inicializa aleatoriamente los vectores α y β , para crear un primer mejor plano (α^* y β^*), luego se genera dos planos vecinos dados por las ecuaciones (2) y se evalúa el índice Chi-cuadrado para ellos, si uno de ellos presenta una mejoría en el índice este será el nuevo mejor plano de lo contrario se generan dos nuevos planos vecinos; si después de cierto número de iteraciones no ha habido mejoría entonces se reduce el tamaño del vecindario de búsqueda a través de la disminución del parámetro c .

$$a_1 = \frac{\alpha^* + cv}{\|\alpha^* + cv\|} \quad b_1 = \frac{\beta^* - (a_1^T \beta^*) a_1}{\|\beta^* - (a_1^T \beta^*) a_1\|}$$

$$a_2 = \frac{\alpha^* - cv}{\|\alpha^* - cv\|} \quad b_2 = \frac{\beta^* - (a_2^T \beta^*) a_2}{\|\beta^* - (a_2^T \beta^*) a_2\|}$$

De forma general el conjunto de pasos para encontrar el mejor índice de proyección son mostrados en la figura 1

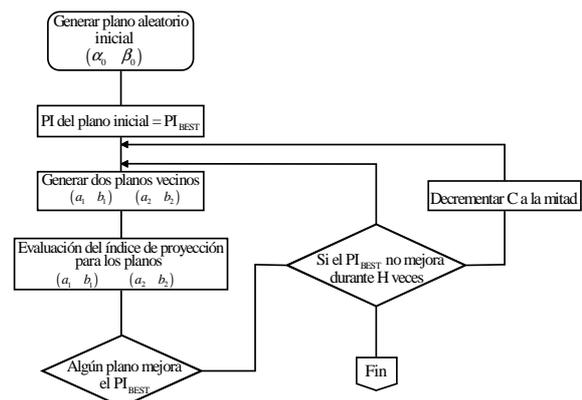


Figura 1. Algoritmo para encontrar el índice de proyección

3.2. Redes neuronales artificiales

Las redes neuronales, como sistemas biológicos, están formados por neuronas de entrada o sensores conectados a una compleja red de neuronas que "calculan", o neuronas ocultas, las cuales, a su vez, están conectadas a las neuronas de salida (como se muestra en la figura 3), encargadas por ejemplo, de controlar los músculos. Por sensores se entienden señales de los sentidos (oído, vista, etc.), las respuestas de las neuronas de salida activan los músculos correspondientes. En el cerebro hay una gigantesca red de neuronas "calculadoras" u ocultas que realizan la computación necesaria. De manera similar, una red neuronal artificial debe ser compuesta por sensores del tipo mecánico o eléctrico.

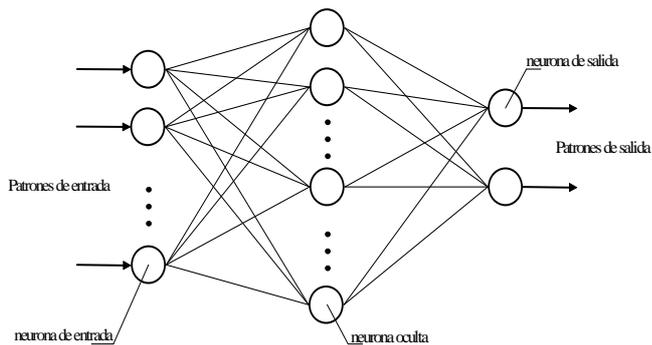


Figura 3. Red Neuronal Artificial Típica

Las Redes Neuronales Artificiales (Artificial Neural Networks) son sistemas paralelos para el procesamiento de la información, inspirados en el modo en el que las redes de neuronas biológicas del cerebro procesan esta. Es decir que se han intentado plasmar los aspectos esenciales de una neurona real a la hora de diseñar una neurona "artificial". Estos modelos realizan una simplificación, desentrañando cuales son las relevancias del sistema.

La definición más general considera a una red neuronal como un entramado o estructura formada por muchos procesadores simples llamados nodos o neuronas, conectados por medio de canales de comunicación o conexiones. Cada una de ellas tiene una cantidad de memoria local, operando solamente con sus datos locales y sobre las entradas que recibe a través de esas conexiones.

Las Redes Neuronales llevan asociadas algún tipo de regla de aprendizaje o entrenamiento particular por la cual esas conexiones son ajustadas acorde a los ejemplos proporcionados. En otras palabras, estas aprenden a partir de ejemplos, y muestran alguna capacidad para generalizar más allá de esos datos mostrados [6.7].

4. METODOLOGÍA PROPUESTA

Lo anteriormente expuesto conlleva que a pesar de no conocerse el modelo que es usado por el Ministerio de Educación, si es determinante en el proceso de calificación de la eficiencia de las universidades y en la posterior distribución de los recursos, los cuales son indispensables para el buen funcionamiento de cada universidad, y como la Universidad Tecnológica tiene el mayor peso en los ingresos anuales y por ende en la capacidad de gestión, una mala ubicación en esta calificación podría llevar al traste con todos los proyectos del Plan de Desarrollo de la Universidad y con su funcionamiento.

Se tiene entonces una valiosa cantidad de información desde el año 2003, no solo de los indicadores que reportan las 32 principales universidades del país y que

contribuyen a la formación del modelo SUE, sino las ubicaciones anuales dadas por el Ministerio de Educación a estas mismas 32 universidades, si se logra obtener un modelo que permita tomar decisiones en función a nuevos datos de entrada, se tendrá la posibilidad de buscar un funcionamiento óptimo para la universidad y poderse colocar a la vanguardia no solo en el cumplimiento de su Plan de Desarrollo sino con respecto a el ranking nacional de las universidades públicas del país.

Para la creación del modelo mencionado es necesario tener en cuenta que el número de variables que representan el conjunto de datos es grande, y que por tal motivo algunas de las variables pueden entregar información redundante que afectará en cierto modo la determinación de los valores de los índices, por tal razón es necesario incluir una técnica de reducción de la dimensionalidad dentro del modelo propuesto. Para la solución del problema de la dimensionalidad (que consiste básicamente en la reducción del espacio dimensional original donde estaban definidas todas las variables originales del problema), se han usado diferentes técnicas [9], y dentro de las más utilizadas gracias a su efectividad en la reducción de bases de datos de gran tamaño, se encuentran, el análisis envolvente de datos (DEA) [7], el análisis de componentes principales (PCA) [4], y la técnica de búsqueda de la proyección (PP) [6]. El éxito de las técnicas mencionadas radica principalmente en que usan técnicas de optimización dentro de sus desarrollos matemáticos (la técnica de mínimos cuadrados en general para DEA y PCA y la maximización del índice Chi-cuadrado para PP), permitiendo encontrar espacios vectoriales que representen de muy buena manera el espacio vectorial original.

Luego de realizar la reducción del espacio de características (paso que se puede tomar como primera etapa del modelo), es necesario tener una segunda etapa en la cual se determinen los valores que debe tener cada uno de los índices mencionados anteriormente. Debido a la característica de los datos del problema y a la efectividad que presentan al solucionar problemas similares, metodologías basadas en la teoría de Bayes (TB), redes neuronales (RN), y maquinas de soporte vectorial (MSV), pueden ser usadas para tal fin [3]. Cada una de las técnicas mencionadas es usada según el estudio que se desea realizar, por ejemplo, los clasificadores o regresores basados en la teoría de Bayes son usados para cuando se quiere probar que tan buena es la caracterización de una determinada base de datos [7], y las RN [12] o MSV[5] son usadas cuando la caracterización de datos no es buena y se desea probar la efectividad de la técnica de clasificación o regresión [3]. Para la conformación del modelo que ayudara a resolver el problema de determinar los valores de los índices de calidad de las universidades, se propone entonces trabajar la primera etapa utilizando la técnica de búsqueda de la proyección (PP), y la segunda con redes neuronales. Una de las razones por la cual se escogen estas dos técnicas es

debida a que ambas usan conceptos de optimización (PP para encontrar una dimensión menor en la cual el conjunto de datos original se encuentre bien representado, y las RN durante el proceso de entrenamiento para minimizar el error y maximizar la capacidad de inferencia frente a la gran cantidad de datos del problema), que de cierta manera garantizan el éxito en sus tareas. Otra de las razones para la elección de las RN, es que para la reducción de la dimensionalidad se usará una técnica fuerte que ha mostrado excelentes resultados en el análisis de bases de datos de gran tamaño [11] [10], y según lo mencionado en el párrafo anterior para probar la efectividad en la reducción, puede ser usada una técnica suave basada en el teorema de Bayes, pero en este caso no se desea realizar dicho análisis, por tal motivo se proponen las RN ya que se encuentran en un punto medio dentro del grupo conformado por TB, RN y MSV.

Es motivo de la investigación entonces el estudio de los indicadores de gestión, el uso de técnicas de reducción de la dimensionalidad en bases de datos de gran tamaño, y de las redes neuronales para la creación de un modelo que permita encontrar los valores más adecuados de los índices para el funcionamiento adecuado de la universidad.

6. BIBLIOGRAFÍA

- [1] Arenas Valencia, Wilson. "Una mirada crítica a los indicadores de desempeño de las Universidades Públicas desde la perspectiva del Análisis Envolvente de Datos-DEA". Tesis de Maestría, año 2005. Universidad Tecnológica de Pereira
- [2] Políticas y estrategias para la educación superior de Colombia 2006 – 2010: "De la Exclusión a la Equidad II. Hacia la construcción de un sistema de educación superior más equitativo y competitivo, al servicio del país", Bogotá, 2007
- [3] C, Burgues, "A tutorial on Classification Techniques for Pattern Recognition, Knowledge Discovery and Data Mining", 2(2), 2003
- [4] Godshalk EB, Timothy DH. "Factor and principal component analyses as alternatives to index selection". *Theoretical Applied Gen* 1998.
- [5] Jian-Pei Zhang, Zhong-Wei Li, Jing Yang, "A parallel SVM training algorithm on large-scale classification problems", *Machine Learning and Cybernetics*, 2005. *Proceedings of 2005 International Conference on Volume 3*, 18-21 Aug. 2005
- [6] L. Jimenez and D. A. Landgrebe, "Projection Pursuit in High Dimensional Data Reduction: Initial Conditions, Feature Selection and the Assumption of Normality", presented at the *IEEE International Conference on Systems, Man and Cybernetics*, Vancouver Canada, October 1995.
- [7] Johnson RA, Wichern DW. "Applied multivariate statistical analysis". U. Wisconsin, Madison. Prentice Inc., 1982
- [8] Ley 30 de 1992, Por la cual se organiza el servicio público de la Educación Superior.
- [9] Montgomery DC, Peck EA. "Introduction to linear regression analysis". 2a Edition, J. Wiley and Sons, New York 1986
- [10] Pérez, L, Mora, J, Pérez, S, "Reducción de las Características asociadas al problema de localización de fallas en sistemas de distribución", *Scientia et Technica*, Vol XIII, N° 35, Agosto de 2007
- [11] Toro, M, Eliana, Pérez, H, Lucas, "Reducción de la dimensionalidad con componentes principales y técnica de búsqueda de la proyección aplicada a la clasificación de nuevos datos", *Revista Tecnura*, Año XIII, No 35, Agosto de 2007. Universidad Distrital Francisco José de Caldas
- [12] Tsukamoto, Y, Namatame, A, "Evolving neural network models", *Evolutionary Computation*, 1996., *Proceedings of IEEE International Conference on 20-22 May 1996*.