

Tipo de artículo: Artículo original
Temática: Inteligencia Artificial
Recibido: 05/01/2015 | Aceptado: 20/02/2015

Evaluación del desempeño de roles en el proceso de formación del Ingeniero Informático

Evaluation of roles in the process of formation of Computer Engineer

Yucely López Trujillo^{1*}, Ingrid Wilford Rivera¹, Margarita André Ampuero¹

¹*Facultad de Informática, Instituto Superior Politécnico “José Antonio Echeverría”, Calle 114 No 11901 e/ Ciclovía y Rotonda, Marianao, La Habana, Cuba CP 19390 {ylopez, iwilford, mayi}@ceis.cujae.edu.cu

* Autor para la correspondencia: ylopez@ceis.cujae.edu.cu

Resumen

En la actualidad el Plan de estudios que se aplica en la carrera de Ingeniería Informática en Cuba, es el denominado Plan D. Su creación y puesta en ejecución busca resolver deficiencias detectadas en los planes de estudios precedentes, en temas como: la creación de habilidades y hábitos de trabajo en equipo y el trabajo por roles del futuro ingeniero, entre otros aspectos. Así, desarrollar un proceso de formación de roles efectivo es un objetivo importante de la carrera. Para el logro de este empeño, resulta imprescindible evaluar sistemáticamente el desempeño de roles por parte de los estudiantes. En el presente trabajo se describe una propuesta para la evaluación del desempeño de roles en la carrera. Por otra parte, para apoyar la toma de decisiones estratégicas vinculadas al proceso de enseñanza, es necesario analizar en profundidad los resultados del proceso de formación de roles, con el propósito de: valorar la calidad de dicho proceso, identificar las fortalezas y debilidades, predecir comportamientos futuros, así como, descubrir relaciones ocultas entre los datos y nuevo conocimiento. Es por esto que se propone la aplicación de la minería de datos.

Palabras clave: calidad, formación, roles, ingeniería informática, minería de datos.

Abstract

At present the curriculum applied in the Software Engineering career in Cuba, is called Plan D. The creation and implementation seeks to address deficiencies identified in previous studies plans on topics such as: building skills and

habits of teamwork and work for the future engineer roles, among others. Thus, developing an effective Roles training process is an important goal. To achieve this commitment, it is essential to systematically assess the performance of roles by students. A proposal for performance evaluation of roles is described in this paper. Moreover, to support strategic decision making related to the educational process, it is necessary to analyze in depth the results of the Training Roles, in order to: assess the quality of the process, identify strengths and weaknesses, predict future behaviors, and discover hidden relationships between data and new knowledge. This is why the application of data mining is proposed.

Keywords: *quality, training, roles, software engineering, data mining.*

Introducción

La carrera de Ingeniería Informática en Cuba (curso regular diurno) ha transitado por diferentes planes de estudio, estructurados en cinco años académicos (diez semestres). Todos los planes han exigido que la totalidad de los estudiantes, a partir del tercer año, se incorporen a la ejecución de proyectos reales, los que en la mayoría de los casos, se desarrollan vinculados a los grupos de investigación de la facultad, donde intervienen estudiantes y profesores. Como parte del proceso de elaboración del plan de estudio vigente (denominado “Plan D”) y en respuesta a los retos que enfrenta el proceso de formación de los profesionales de software, se elaboró una estrategia para la formación de los roles y las competencias que precisa la industria en cuanto a trabajo disciplinado y en equipo, la que incorpora, entre otros aspectos, la introducción de las prácticas de PSP y TSP de forma incremental (ANDRE, 2007) (LOPEZ, 2011).

En la actualidad, el claustro de profesores de la carrera de Ingeniería Informática a nivel nacional, no realiza la evaluación del desempeño de roles de manera formalizada y por tanto no se realizan análisis de dichos resultados, lo que afecta la calidad del proceso de formación. Para perfeccionar continuamente la formación de ingenieros informáticos, tomar decisiones estratégicas y proponer mejoras, es necesario analizar en profundidad los resultados de las evaluaciones del desempeño realizadas a los estudiantes durante su formación de pregrado.

Sin embargo, debido al volumen de dicha información y al poco tiempo del que disponen los profesores de la carrera, no es viable procesarla por los métodos tradicionales, ya que el análisis podría ser muy superficial. En este sentido, puede resultar muy útil la aplicación de la minería de datos, una generación de herramientas y técnicas que soportan la extracción de conocimiento útil, a partir de la información disponible. (CIOS, 2009) (WITTEN, 2005), la aplicación

de la minería de datos está presente en múltiples entornos (ROMERO, 2010) (SREEVIDYA, 2014) (WESLEY, 2013) (FOSTER, 2013) (ZAKI, 2014). Es por ello que este trabajo propone el uso de la minería de datos para el análisis de la calidad del proceso de formación de roles en la carrera, a partir de las variables que miden las habilidades técnicas y genéricas desarrolladas por los estudiantes.

El presente trabajo se estructura como sigue: se realiza un análisis del proceso de formación de roles que propone el Plan D, se enuncia una propuesta de evaluación del desempeño de roles para ser utilizada por cada una de las asignaturas que tributan a este empeño y por último, se propone el uso de la minería de datos para el análisis de la calidad del proceso de formación de roles.

Materiales y métodos.

Resulta necesario comenzar este acápite con un análisis del proceso de formación de roles en la carrera de Ingeniería Informática, teniendo en cuenta que este proceso constituye el material de partida para enunciar la elección de métodos y técnicas presentes en la propuesta de evaluación del desempeño de roles.

Proceso de formación de roles en la carrera de Ingeniería Informática

El Plan de estudios vigente, Plan D, que se aplica en la Facultad de Informática de la CUJAE, tuvo sus antecedentes en los análisis críticos realizados a su predecesor, el denominado Plan C. Así en (ANDRE, 2007), se realiza un análisis crítico del proceso de formación del ingeniero informático con el Plan C y se concluye que los temas de planificación, asignación de recursos, estimación de costos, tamaños y tiempos, definición y trabajo con estándares, calidad, entre otros no eran abordados de forma profunda.

Con el objetivo de mejorar el proceso de formación concebido en el Plan C se propuso una estrategia, incluida en el Plan D, para lograr no sólo que los estudiantes adquirieran una disciplina personal, sino que aprendieran a desempeñar los roles a lo largo de toda la carrera de forma incremental, consciente y disciplinada con vistas a poder ser miembros efectivos de un equipo de proyecto. En la concepción de la estrategia de formación de roles, se toma en cuenta que la formación es uno de los procesos claves de la Gestión de los Recursos Humanos, (PMI, 2013) (ANDRE, 2010) (ANDRE, 2011) (HUSSEY, 2011) (INFANTE, 2011). En el ámbito de la enseñanza de la carrera de ingeniería informática, resulta importante concebir al “Proceso de Formación” en su sentido completo, no únicamente como la actividad de instruir para un determinado rol, sino como la formación de conocimientos y habilidades.

En la estrategia se identifican doce roles a formar durante la carrera: Analista (negocio y sistema), Diseñador, Programador, Arquitecto, Planificador, Especialista en calidad, Especialista en seguridad, Jefe de proyecto, Implantador, Probador, Especialista en soporte y Gestor de configuración y cambios. De ellos, son secundarios: el Especialista en Seguridad, el Implantador y el Especialista en Soporte; mientras que los restantes son primarios en al menos un año académico. Se proponen entonces como roles primarios en la formación de un ingeniero informático: el Analista, el Arquitecto, el Diseñador, el Programador, el Jefe de proyecto, el Especialista en calidad, el Planificador, el Probador y el Gestor de configuración y cambios. El Proceso de Formación para la totalidad de los roles es iterativo e incremental, alcanzándose en el último año de la carrera el 100% de la formación requerida en cada uno de los casos. Resulta válido señalar que en la propuesta (ANDRE, 2007) se encuentra un análisis más detallado de cómo tributa cada asignatura a la formación de los roles en cada año académico precisando por cada una los objetivos fundamentales, los roles primarios y secundarios y las tareas definidas para cada rol.

A partir del curso académico 2007-2008 se comenzó a aplicar el Plan D, en la facultad de Ingeniería Informática del Instituto Superior Politécnico José Antonio Echeverría. El plan de estudio D tiene una organización docente formada por un currículo base, un currículo propio y un conjunto de asignaturas optativas. Dentro de las asignaturas del currículo propio, de la disciplina “Ingeniería y Gestión de Software” que tributan directamente a la formación de roles y al trabajo en equipo se encuentra la asignatura Ingeniería de Software III (IS3), en estos momentos se cuenta con una experiencia en su ejecución de cuatro cursos académicos.

En (LOPEZ, 2011) se realiza un análisis más detallado de cómo se estructura y se lleva a la práctica la asignatura IS3. Ahora bien, partiendo de la experiencia que se tiene en la impartición de la asignatura IS3, se puede constatar que los profesores pertenecientes a los colectivos de asignaturas que tributan a la formación de roles no están evaluando el desempeño de estos formalmente. Al no evaluar los roles los estudiantes no sienten que están siendo formados para su desempeño de forma consciente. Este aspecto, hace que en la asignatura IS3 haya que asignar los roles, auxiliándose de la aplicación de diversos instrumentos que buscan descubrir si los estudiantes tienen las competencias tanto, técnicas como genéricas. En ocasiones, los roles, para la conformación del equipo se asignan de acuerdo al buen juicio de los miembros del equipo sin tener criterios reales.

Es por esta razón que urge en la carrera de Ingeniería informática contar con una propuesta que permita evaluar el desempeño de roles de manera formalizada. Esta propuesta debe proveer a los colectivos de profesores, de instrumentos que permitan que se evalúe de manera uniforme y consciente. Por otra parte, en esta propuesta debe

concebirse la evaluación de las habilidades no solo técnicas, sino también genéricas, entendiéndose por habilidades genéricas aquellas que no están asociadas al desarrollo de una profesión en específico, como por ejemplo las habilidades para la comunicación. A continuación se enuncia la propuesta de evaluación del desempeño de roles, que tiene sus antecedentes en (WILFORD, 2006).

Propuesta de evaluación del desempeño de roles

A través de la Evaluación del Desempeño es posible determinar las habilidades técnicas y genéricas que han desarrollado los estudiantes de la carrera de Ingeniería Informática, identificando las fortalezas y debilidades del proceso de formación de roles, de lo que se derivan las necesidades de formación.

Para la evaluación de las habilidades técnicas se propone:

- Evaluarlas en cada uno de los cortes evaluativos para las asignaturas del plan de estudio que tributan a la formación de roles.
- Evaluarlas a través de la calidad de los artefactos creados o actualizados por los estudiantes.
- Para cada artefacto a evaluar es necesario precisar el rol o los roles responsables de su creación o actualización, ya que en un mismo corte evaluativo es posible evaluar el desempeño de más de un rol.
- Para evaluar la calidad de los artefactos, se emplearán Formularios de Evaluación que se deberán elaborar de acuerdo a las Listas de Chequeo.
- Las Listas de Chequeo deben ser previamente seleccionadas y definidas. Además, deben ser publicadas entre los estudiantes para que sea efectivo un proceso de revisión y autoevaluación previo.
- Para cada rol a evaluar en un corte evaluativo de una determinada asignatura, se deberá elaborar un Formulario de Evaluación.
- La calificación final de un estudiante en un corte evaluativo de una asignatura dada, deberá depender de la calidad del desempeño de los diferentes roles involucrados.
- Cada Formulario de Evaluación, correspondiente a un rol determinado, deberá incluir un fragmento para la evaluación de cada artefacto requerido. Estos fragmentos deben estar formados por preguntas correspondientes a las frases incluidas en las Listas de Chequeo de los artefactos que serán evaluados. Es necesario definir para cada pregunta, las posibles variantes de respuestas, por ejemplo: (Sí, No) o (Alta, Media, Baja, Muy Baja), entre otras.

- Para cada una de las preguntas que componen los fragmentos de evaluación de los artefactos, y según la respuesta, se deberá definir un puntaje o peso. El valor mayor de peso para una pregunta, se deberá asignar a la respuesta correcta o deseada.
- Cada artefacto, tiene un puntaje máximo, que se calcula sumando los puntajes correspondientes a las respuestas correctas o deseadas.
- Para cada artefacto a evaluar es necesario definir rangos de porcentajes que se correspondan con los posibles resultados o valoraciones relativas a su calidad. Estas valoraciones podrían ser: Alta, Media, Baja, Muy Baja.
- Además, para cada artefacto a evaluar en un Formulario de Evaluación, es necesario definir su porcentaje de significación máxima. Esto indica qué importancia tiene o cuánto pesa su calidad en la valoración final del desempeño del rol correspondiente.
- Para cada Formulario de Evaluación, es necesario definir rangos de porcentajes que se correspondan con los posibles resultados finales sobre el desempeño del rol que se evalúa: Óptimo, Bueno, Regular y Malo, en correspondencia con el sistema de evaluación definido en la universidad de 5, 4, 3 y 2, respectivamente.
- Para determinar la calificación final de un estudiante en un corte evaluativo de una asignatura dada, es necesario definir, para cada Formulario de Evaluación que evalúa un determinado rol, su porcentaje de significación máxima. Esto indica qué importancia tiene o cuánto pesa el desempeño de dicho rol en la calificación final del corte evaluativo.

Para la evaluación de las habilidades genéricas se propone:

- Evaluar las habilidades genéricas al concluir la asignatura.
- Para evaluar las habilidades genéricas desarrolladas por los estudiantes durante su formación, es conveniente considerar, tanto el criterio de los evaluadores, como la valoración que cada estudiante tiene de sí mismo. Para ello, es necesario elaborar Tests de Evaluación y Autoevaluación de dichas habilidades.
- Evaluar una habilidad genérica determinada implica:
 - Evaluar un conjunto de habilidades derivadas de la habilidad genérica. Por ejemplo, para la habilidad genérica “habilidades para la comunicación”, evaluar las habilidades derivadas: habilidades para la expresión escrita, habilidades para la expresión oral, habilidades para la escucha y habilidades para la observación.

- Evaluar elementos de carácter metacognitivo. Estos propician un constante perfeccionamiento del desempeño del individuo y se refieren a los siguientes indicadores o conductas no observables.
 - Conoce sus límites y potencialidades.
 - Reflexiona acerca de estos aspectos.
 - Pone en práctica estrategias de superación de dificultades.
 - Existe congruencia entre la imagen que tiene de sí y la que proyecta.
- Para evaluar los elementos de carácter metacognitivo, en cada Test de Autoevaluación correspondiente a una habilidad genérica determinada, se enunciarán los indicadores o conductas no observables que los detallan, de los que se deberá seleccionar la respuesta adecuada: (Sí, No, No Sé).
- Para evaluar las habilidades o características derivadas de una habilidad genérica determinada, en los Tests de Evaluación y Autoevaluación correspondientes, cada habilidad o característica derivada deberá estar descrita a través de un conjunto de indicadores o conductas de las que se deberá seleccionar la frecuencia en que tienen lugar: (Casi Siempre, A Veces, Casi Nunca).
- Los enunciados de los indicadores o conductas deberán ser concretos y claros, redactándose en tercera persona en el caso de los Test de Evaluación y en primera persona en el caso de los de Autoevaluación.
- Para cada indicador o conducta de una habilidad genérica es preciso definir un puntaje o peso, donde el valor mayor se corresponde con la respuesta deseada.
- Así, cada habilidad o característica derivada que permite evaluar una habilidad genérica dada, tiene un puntaje máximo, que se calcula sumando los puntajes correspondientes a las respuestas deseadas de cada uno de los indicadores o conductas que los describen. Para cada habilidad genérica y para las habilidades o características derivadas, es preciso definir rangos de porcentajes de “calidad”, que se correspondan con las posibles valoraciones resultantes de procesar los Tests de Evaluación y Autoevaluación aplicados. Por ejemplo, para las siguientes valoraciones: “Mala”, “Regular”, “Buena” y “Muy Buena”, es posible definir los rangos de porcentajes de “calidad”: [0%...60%], [60%...70%], [70%...90%] y [90%...100%] respectivamente.
- Para cada una de las habilidades o características derivadas y para el conjunto de componentes metacognitivos relativos a una habilidad genérica dada, es preciso definir su porcentaje de significación máxima. Esto indica la importancia que tienen o cuánto pesan en la valoración final de la habilidad genérica correspondiente. La sumatoria de todos estos porcentajes, para un test determinado, debe ser igual a 100%. Por ejemplo, para la

habilidad genérica “habilidades para la comunicación”, las habilidades derivadas podrían tener los siguientes porcentajes: habilidades para la expresión oral (30%), habilidades para la expresión escrita (30%), habilidades para la escucha (30%), habilidades para la observación (10%). Sin embargo, en el Test de Autoevaluación de la misma habilidad genérica, estos porcentajes deben variar, ya que se evalúa, además, el conjunto de componentes metacognitivos, para el que se debe definir, también, su porcentaje de significación máxima.

Después de definir o seleccionar los formularios de evaluación para las habilidades técnicas, y los tests de evaluación y autoevaluación para las habilidades genéricas, los evaluadores deben registrar los resultados y procesarlos, mediante la herramienta EvalSoft¹, para poder realizar los análisis necesarios. Sin embargo, debido al volumen de dicha información, no es viable procesarla por los métodos tradicionales. El análisis e interpretación manual de esta información resulta sumamente lento y podría ser muy superficial. Es por ello que se propone utilizar técnicas de minería de datos, ya que estas enriquecen el proceso de análisis, descubriendo relaciones ocultas entre los datos e información desconocida, favoreciendo la gestión y toma de decisiones en los procesos del negocio.

Resultados y discusión

Como resultado de la Evaluación del Desempeño de roles, actividad que se deberá realizar de forma sistemática, es posible obtener numerosos y valiosos indicadores, entre estos se encuentran: las variables que miden la calidad de los artefactos elaborados (como resultado de la evaluación de las habilidades técnicas) y las variables que miden las habilidades genéricas desarrolladas por los estudiantes durante su formación. Estos indicadores deberán ser analizados en profundidad con el propósito de:

- Identificar fortalezas y debilidades del proceso de formación de ingenieros informáticos.
- Obtener información útil y nuevo conocimiento, que apoye la toma de decisiones vinculadas al proceso de enseñanza de la carrera. (Identificar relaciones ocultas entre los datos.)
- Validar el proceso de evaluación que siguen las diferentes asignaturas de la carrera que tributan a la formación de roles. (Ajustar pesos, puntajes, entre otros (WILFORD, 2006)).
- Realizar propuestas de mejoras continuas, vinculadas al proceso de formación de roles en la carrera.

La minería de datos, enriquece el proceso de análisis, descubriendo relaciones ocultas entre los datos e información desconocida; por lo que se propone la aplicación de técnicas de minería de datos para el análisis de dichos

¹ EvalSoft es una herramienta para la evaluación de roles, resultado del proceso de investigación y desarrollo propuesto en (WILFORD, 2006).

indicadores. Para ello, es necesario realizarlas siguientes tareas: definir las variables a utilizar en la minería de datos, definir la variable objetivo y seleccionar las técnicas de minería de datos a ser empleadas.

Aplicación de la Minería de Datos: definición de las variables para la minería de datos

Previo a la definición de las variables de entrada para la aplicación de las técnicas de minería de datos, es conveniente detallar los objetivos que se persiguen con el análisis de los datos disponibles. Estos objetivos son los siguientes:

1. Relacionar las variables que miden la calidad general de los artefactos elaborados por los estudiantes, con las variables que miden la calidad del desempeño del rol que se evalúa. (Los valores de las variables que miden la calidad general de los artefactos y la calidad del desempeño del rol que se evalúa, se obtienen a partir de procesar los formularios de evaluación de las habilidades técnicas correspondientes (WILFORD, 2006)).
2. Relacionar las variables que constituyen indicadores de la calidad de un determinado artefacto, con las variables que miden la calidad general de dicho artefacto y las variables que miden la calidad del desempeño del rol que se evalúa.
3. Relacionar las variables que miden las habilidades genéricas desarrolladas por los estudiantes, con las variables que miden la calidad del desempeño del rol que se evalúa. (Los valores de las variables que miden las habilidades genéricas desarrolladas por los estudiantes, se obtienen a partir de procesar los tests de evaluación y autoevaluación de las habilidades genéricas correspondientes (WILFORD, 2006)).
4. Relacionar las variables que miden las habilidades derivadas desarrolladas por los estudiantes, con las variables que miden las habilidades genéricas y las variables que miden la calidad del desempeño del rol que se evalúa. (Los valores de las variables que miden las habilidades derivadas desarrolladas por los estudiantes, se obtienen a partir de procesar los tests de evaluación y autoevaluación de las habilidades derivadas correspondientes (WILFORD, 2006)).
5. Relacionar las variables que miden las conductas que describen una determinada habilidad derivada, con las variables que miden las habilidades derivadas correspondientes, y las variables que miden la calidad del desempeño del rol que se evalúa.
6. Relacionar las variables que miden la calidad del desempeño de los roles evaluados en un “corte evaluativo” determinado, con las variables que miden el resultado final de la evaluación de las habilidades técnicas requeridas en dicho “corte evaluativo”.

A partir de los objetivos planteados, es posible definir los siguientes grupos de variables a utilizar en la minería de datos:

1. Datos generales de los evaluados, por ejemplo: “identificador”, “grupo docente”, “vía de ingreso”, “índice académico de ingreso”, entre otros que resulten de interés.
2. Variables que miden el resultado final de la evaluación de las habilidades técnicas requeridas en un “corte evaluativo” dado. En este grupo se incluyen las variables que almacenan la *Valoración Final* sobre la calidad general del desempeño de todos los roles evaluados en dicho “corte evaluativo”, cuyos valores se obtienen a partir del procesamiento de los formularios de evaluación de las habilidades técnicas correspondientes. Además, se incluye en este grupo, la variable que almacena la *Calificación Final* correspondiente al “corte evaluativo”, determinada por los evaluadores de las habilidades técnicas.
3. Variables que miden la calidad del desempeño de un determinado rol. En este grupo se incluyen las variables que almacenan la *Valoración Final* sobre la calidad del desempeño de un rol dado, cuyos valores se obtienen a partir del procesamiento del formulario de evaluación de las habilidades técnicas correspondientes. Por ejemplo, una variable puede ser “*DESEMPEÑO_Analista_Negocio_Valoración Final*” (*Valoración Final* de la calidad del desempeño del rol Analista del Negocio), donde los valores posibles que dicha variable puede tomar podrían ser: “Óptimo”, “Bueno”, “Regular” o “Malo”).
4. Variables que miden la calidad general de los artefactos elaborados. En este grupo se incluyen las variables que almacenan la *Valoración Final* sobre la calidad de los artefactos evaluados, cuyos valores se obtienen a partir del procesamiento del formulario de evaluación de las habilidades técnicas correspondientes.
5. Variables que constituyen indicadores de la calidad de un determinado artefacto. Estas variables se corresponden con cada una de las frases que conforman la lista de chequeo relativa al artefacto evaluado y que se incluyen en los formularios de evaluación.
6. Variables de coincidencia de habilidades genéricas. Las variables incluidas en este grupo indican si “hay coincidencia” o no entre los resultados de los test de autoevaluación y evaluación para una habilidad genérica determinada.
7. Variables de coincidencia de habilidades derivadas. Las variables incluidas en este grupo indican si “hay coincidencia” o no entre los resultados de los test de autoevaluación y evaluación para una habilidad derivada de una habilidad genérica determinada.
8. Variables que miden las habilidades genéricas desarrolladas por los estudiantes. Estas variables se corresponden con las *Valoraciones Finales* sobre las habilidades genéricas desarrolladas por los estudiantes,

obtenidas a partir del procesamiento de los test de evaluación y autoevaluación correspondientes. Por ejemplo, una variable podría referirse a la *Valoración Final* sobre las habilidades para la comunicación desarrolladas, según resultados de los test de evaluación que han sido procesados, donde los valores posibles que dicha variable puede tomar podrían ser: “Malas”, “Regulares”, “Buenas” o “Muy Buenas”.

9. Variables que miden las habilidades derivadas desarrolladas por los estudiantes. Estas variables se corresponden con las *Valoraciones Finales* sobre las habilidades derivadas desarrolladas por los estudiantes, obtenidas al procesar los test de evaluación y autoevaluación relativos a la habilidad genérica correspondiente.
10. Variables que miden las conductas que describen las diferentes habilidades derivadas. Estas variables se corresponden con las valoraciones referentes a las conductas que describen las diferentes habilidades derivadas, evaluadas en los test de evaluación y autoevaluación relativos a la habilidad genérica correspondiente. Un ejemplo de variable que se puede incluir en este grupo podría ser la que se refiere a la conducta: “*Utiliza un lenguaje directo, sin rodeos*”, donde los valores posibles que dicha variable puede tomar podrían ser: “Casi Siempre”, “A Veces” o “Casi Nunca”.

En correspondencia con los objetivos planteados anteriormente, se define “DESEMPEÑO” como variable objetivo, a partir de la Valoración Final y la Calificación Final:

- *Valoración Final* resultante del procesamiento de los formularios de evaluación. (Se refiere a la calidad del desempeño de un rol dado, en un “corte evaluativo” de una asignatura determinada, o a la calidad general del desempeño de todos los roles implicados en dicho “corte evaluativo”, según el objetivo.)
- *Calificación Final* determinada por los evaluadores de las habilidades técnicas. (Se refiere a la calidad del desempeño de un rol dado, en un “corte evaluativo” de una asignatura determinada, o a la calidad general del desempeño de todos los roles implicados en dicho “corte evaluativo”, según el objetivo.)

La variable objetivo DESEMPEÑO Valoración Final, según la propuesta, toma valores nominales (Óptimo, Bueno, Regular, Malo); mientras que, la variable objetivo DESEMPEÑO Calificación Final, toma valores numéricos (5, 4, 3, 2), por lo que, en este caso, se decidió categorizar dicha variable como se muestra en la Tabla 1:

Tabla1. Clases de DESEMPEÑO.

Calificación Final	Clases de Desempeño
2	Malo

3	Regular
4	Bueno
5	Óptimo

Una vez definidos los objetivos y las variables necesarias, es preciso seleccionar las técnicas y algoritmos de minería de datos a emplear.

Selección de técnicas de Minería de Datos y herramientas

Para cumplir los objetivos antes enunciados se propone la aplicación conjunta de las siguientes técnicas de minería de datos: Clustering o agrupamiento, Determinar grupos afines o reglas de asociación y Clasificación (HAN, 2006) (HERNANDEZ, 2004) (WITTEN, 2005) (SUGANTHI, 2014) (LIAO, 2012) SREEVIDYA, 2014) (WESLEY, 2013) (FOSTER, 2013) (ZAKI, 2014).

- Clasificación. Identificación de características de un objeto o registro con el propósito de asignarle una clase o categoría predefinida. Para ello, se requiere construir un modelo de clasificación. La salida obtenida son valores discretos, que se distribuyen en grupos o clases. Para la clasificación existen varios tipos de técnicas: métodos de inducción de reglas, árboles de decisión, redes neuronales, algoritmos tipo k-nn (k-nearest neighbours), métodos bayesianos, entre otros.
- Determinación de grupos afines o reglas de asociación. Se encarga de descubrir fenómenos que ocurren de conjunto, aunque se desconoce el tipo de relación causal que existe entre estos. A partir de los grupos afines identificados es posible, generar reglas de asociación entre los datos. Una regla de asociación constituye una implicación $X \rightarrow Y$, en la que X (antecedente) y Y (consecuente) representan conjuntos de pares atributo-valor. Si un atributo determinado aparece en el antecedente de una regla, entonces no aparecerá en el consecuente de la misma, y viceversa. Uno de los algoritmos más populares para generar reglas de asociación, y en el que se basan otros muchos algoritmos, es el Apriori.
- Agrupamiento o Clustering. Tiene el propósito de formar subgrupos homogéneos (clusters), a partir de un grupo diverso, según el grado de semejanza entre las instancias; los elementos de un cluster tienen una “similitud” alta entre ellos y baja con respecto a los elementos de otros clusters. La formalización del concepto de “similitud” es a través de métricas o medidas de distancia. Para implementar esta tarea se han desarrollado diferentes técnicas: métodos aglomerativos jerárquicos, divisivos jerárquicos, particionales, probabilísticos, entre otros.

Para la implementación de la propuesta de análisis de la calidad del proceso de formación de roles en la carrera de ingeniería informática, se propone un sistema que integra las herramientas: EvalSoft y KNIME (*Konstanz Information Miner*).

EvalSoft se encarga de automatizar las fases de recopilación, preparación y pre-procesamiento de los datos, con el objetivo de proporcionarle a una herramienta de minería de datos, las fuentes de datos necesarias para la aplicación de las técnicas de minería de datos seleccionadas (WILFORD, 2006).

KNIME es una plataforma de exploración de datos modular que guía el proceso de descubrimiento de conocimiento en bases de datos. Fue desarrollada originalmente en el departamento de bioinformática y minería de datos de la Universidad de Constanza, Alemania. Es una herramienta de código abierto, multiplataforma, y compuesta por módulos ensamblables que permiten extenderla. Se basa en el diseño de flujos de trabajo (*workflows*) de forma visual, y permite además ejecutar selectivamente algunos o todos los pasos de análisis, para luego investigar los resultados a través de vistas interactivas de los datos y modelos (BERTHOLD, 2007).

El sistema propuesto permitirá identificar las fortalezas y debilidades de la carrera en relación al proceso de formación de los roles que precisa la Industria Nacional de Software, así como descubrir posibles relaciones entre las variables definidas, ayudando a proponer mejoras con el objetivo de perfeccionar la enseñanza de la ingeniería informática en Cuba.

Conclusiones

- Es preciso modificar la forma en que tiene lugar, en la actualidad, el proceso de Evaluación del desempeño de roles en la carrera de Ingeniería Informática, para ello resulta conveniente poner en práctica la propuesta de evaluación enunciada en este trabajo.
- La propuesta se basa en la evaluación de las habilidades técnicas y genéricas desarrolladas por los estudiantes durante su formación, y en la aplicación de técnicas de minería de datos para el análisis de dichos resultados.
- Para la implementación, se propone un sistema que integra las herramientas: EvalSoft y KNIME. EvalSoft para proporcionarle a la plataforma KNIME las fuentes de datos necesarias para la aplicación de las técnicas de minería de datos seleccionadas.
- La propuesta permitirá identificar las fortalezas y debilidades de la carrera en relación al proceso de formación de los roles que precisa la Industria Nacional de Software y brindará elementos para proponer mejoras con el objetivo

de perfeccionar la enseñanza de la Ingeniería Informática en Cuba.

Actualmente se trabaja en la actualización de las listas de chequeo a utilizar para evaluar la calidad de los artefactos creados por los estudiantes en cada corte evaluativo de las asignaturas. Una vez que se disponga de las listas de chequeo actualizadas se procederá a aplicar la propuesta mediante un caso de estudio. En trabajos futuros se pretende analizar resultados experimentales que permitan valorar la aplicabilidad del sistema propuesto y de las técnicas de minería de datos seleccionadas.

Referencias

ANDRE AMPUERO, M., *Estrategia para la formación de roles en la carrera de Ingeniería Informática basado en la gestión por competencias, Parte 1*, in *Reporte de Investigaciones del Centro de Estudio de* AGRAWAL, R., IMIELINSKI, T., SWAMI, A. Mining association rules between sets of items in large databases. In: Proceedings of the ACM SIGMOD International Conference on Management of Data. 1993. Washington D.C., p. 207-216.

BROWN, Meta S. *Data mining for dummies*. John Wiley & Sons, 2014.

BRIN, S., MOTWANI, R., ULLMAN, J. D., TSUR, S. Dynamic itemset counting and implication rules for market basket data. Proceedings ACM SIGMOD International Conference on Management of Data. 1997. Tucson, Arizona, USA. p. 255-264

CHEN, M., HAN, J. AND YU P. Data Mining: An Overview from a Database Perspective. IEEE Transactions on Knowledge and Data Engineering, 1996. V.8 N.6, p.866-883.

CONEAU- <http://www.coneau.gov.ar/>

HAHSLER, Michael; HORNIK, Kurt; REUTTERER, Thomas. Implications of probabilistic data modeling for mining association rules. En *From Data and Information Analysis to Knowledge Engineering*. Springer Berlin Heidelberg, 2006. p. 598-605.

HERRERA VARELA, Ricardo. Bibliomining: minería de datos y descubrimiento de conocimiento en bases de datos aplicados al ámbito bibliotecario [en línea]. "Forinf@", vol. 33, 2006. [Consulta: 01/09/2014].

<<http://lemi.uc3m.es/est/forinf@/index.php/Forinfa/rt/printerFriendly/122/127>>

HIPP, Jochen; GÜNTZER, Ulrich; NAKHAEIZADEH, Gholamreza. Algorithms for association rule mining—a general survey and comparison. *ACM sigkdd explorations newsletter*, 2000, vol. 2, no 1, p. 58-64.

KANTARDZIC, M. *Data Mining: Concepts, Models, Methods, and Algorithm*. John Wiley & Sons .2003.343 pages.

LENCA, P., MEYER, P., VAILLANT, B., LALLICH, S. On selecting interestingness measures for association rules: user oriented description and multiple criteria decision aid. [en línea]. *European Journal of Operational Research*. Volume 184, Issue 2, 16 January 2008, Pages 610–626. [Consulta: 01/03/2008]. <http://www.sciencedirect.com/science/article/pii/S0377221706011465>

NICHOLSON, S. The basis for bibliomining: Frameworks for bringing together usage-based data mining and bibliometrics through data warehousing in digital library services. [en línea]. *Information Processing & Management*. Volume 42, Issue 3, May 2006, Pages 785–804 [Consulta: 01/02/2008]. <http://www.sciencedirect.com/science/article/pii/S0306457305000658>

LÓPEZ, Ana Pérez. *La evaluación de colecciones: métodos y modelos*. *Documentación de las Ciencias de la Información*, 2002, vol. 25, p. 321-360.

LUCAS, Joel Pinho. *Métodos de clasificación basados en asociación aplicados a sistemas de recomendación*. 2010. Tesis Doctoral. Universidad de Salamanca.

ROMERO, C. *Aplicación de técnicas de adquisición de conocimiento para la mejora de cursos hipermedia adaptativos basados en Web*. Tesis Doctoral.. Universidad de Granada. E.T.S. Ingeniería Informática. 2003.

SAHU, Hemlata; SHRMA, Shalini; GONDHALAKAR, Seema. A Brief Overview on Data Mining Survey. *International Journal of Computer Technology and Electronics Engineering (IJCTEE) Volume*, 2011, vol.1.

SILVERSTEIN, Craig; BRIN, Sergey; MOTWANI, Rajeev. Beyond market baskets: Generalizing association rules to dependence rules. *Data mining and knowledge discovery*, 1998, vol. 2, no 1, p. 39-68.

UNIVERSIDAD DE BUENOS AIRES. SISTEMA DE BIBLIOTECAS Y DE INFORMACIÓN. COMISIÓN TÉCNICA DE ESTÁNDARES. *Estándares del Sistema de Bibliotecas de la Universidad de Buenos Aires*. SISBI, Universidad de Buenos Aires, Secretaría de Ciencia y Técnica, Sistema de Bibliotecas y de Información, 2013.

VAILLANT, Benoît; LENCA, Philippe; LALLICH, Stéphane. A clustering of interestingness measures. En *Discovery Science*. Springer Berlin Heidelberg, 2004. p. 290-297.

ZHAO, Yanchang; ZHANG, Chengqi; ZHANG, Shichao. Discovering interesting association rules by clustering. En *AI 2004: Advances in Artificial Intelligence*. Springer Berlin Heidelberg, 2005. p. 1055-1061.