

Detección de los niveles de madurez de la gestión del conocimiento en el aprendizaje de un curso de programación de ordenadores

Luis Alberto Muñoz Gómez, luis.munoz@academicos.udg.mx
María Elena Romero Gastelú, elena.romero@academicos.udg.mx
Janette Araceli Castellanos Barajas, janette.castellanos@academicos.udg.mx
Abelardo Gómez Andrade, abelardo.gandrade@academicos.udg.mx
Juan Carlos González Aldana, juan.gonzalez@alumnos.udg.mx

Universidad de Guadalajara, Centro Universitario de Ciencias Exactas e Ingenierías, Departamento de Ciencias Computacionales,
Blvd. Marcelino García Barragán #1421, esq. Calzada Olímpica, C.P. 44430, Guadalajara, Jalisco, México.

Resumen

La presente investigación tiene como objeto determinar cuál es el nivel de madurez de la gestión del conocimiento de un estudiante, mismo que puede ser inicial, conciencia, definido, gestionado u optimizado, basados estos en el Modelo llamado General Knowledge Management Maturity Model (GKMMM) y que inciden en el aprendizaje de un curso introductorio de programación de ordenadores, de las carreras de Ingeniería en Computación, Ingeniería Informática, Ingeniería Biomédica, Ingeniería Electrónica e Ingeniería Civil de los ciclos escolares 2020A y 2020B, para lo cual se implementó y aplicó una encuesta a alumnos, cuyos datos se analizaron e interpretaron de forma automatizada con un sistema diseñado para este fin y que fue denominado "Sistema de Detección de Madurez para Aprender Programación".

Palabras clave

Ingeniería, Gestión del conocimiento, GKMMM, Madurez, Programación de ordenadores.

Abstract

The present research aims to determine what is maturity level of a student's knowledge management, which can be initial, repeatable, defined, managed or optimizing, based on the Model called General Knowledge Management Maturity Model (GKMMM) and that affect the learning of an introductory course in computer programming, of the careers of Computer Engineering, Informatics, Biomedical Engineering, Electronic Engineering and Civil Engineering of the school cycles 2020A and 2020B, for which a survey of students, whose data were analyzed and interpreted in an automated way with a system designed for this purpose and which was called "System of Detection of Maturity to Learn Programming".

Keywords

Computer programming, Engineering, GKMMM, Knowledge management, Maturity.

Antecedentes

En el Centro Universitario de Ciencias Exactas e Ingenierías (CUCEI) de la Universidad de Guadalajara, no ha existido un proyecto que se encargue de determinar los factores que inciden en el aprovechamiento de los estudiantes para un curso en particular, motivo por el cual se optó por realizar el análisis del mismo.

En la Universidad de Guadalajara, hasta antes del inicio de la pandemia se aplicaba un modelo de enseñanza-aprendizaje presencial. El ciclo escolar 2020A inició de forma presencial, pero se vio interrumpido debido a la pandemia a partir del 13 de marzo de 2020, por lo que se concluyó el mismo en una modalidad híbrida; una situación particular en la que las autoridades de la institución decidieron que no quedarán registradas las calificaciones que no fueran aprobatorias, en apoyo a los estudiantes por sus condiciones de conectividad o económicas. Por otro lado, el ciclo 2020B se impartió por completo en una modalidad híbrida. Originalmente se tenía contemplado evaluar la madurez y el desempeño de los alumnos en condiciones normales (presenciales), y en estos dos ciclos atípicos, al modificarse la forma de la enseñanza y el cómo los alumnos adquirirían los conocimientos, esto pudiese haber influido en los resultados esperados de este estudio, debido al cambio en el proceso de enseñanza-aprendizaje en ambos ciclos escolares.

Introducción

Justificación

Desde hace varios ciclos escolares, la mayoría de los autores de esta investigación han impartido el curso introductorio de programación, asignatura que es fundamental para cumplir con el perfil de egreso de estudiantes de las áreas de ciencias exactas e ingenierías, y por ello ya habían tenido la intención de llevar a cabo estrategias para mejorar el aprovechamiento de sus estudiantes, y ahora con este proyecto de investigación se concreta esta intención y se hace un estudio formal para lograr este objetivo.

Descripción del problema

Para toda unidad de aprendizaje existe un índice de aprovechamiento. La unidad de aprendizaje de "Programación", para toda carrera orientada a las TICs (Tecnologías de la Información y la Comunicación) de nivel superior, es fundamental para la adecuada adquisición de conocimientos en muchas otras unidades de aprendizaje posteriores de la carrera, como por ejemplo "Estructuras de Datos". El nivel de aprovechamiento logrado en "Programación" incide de manera importante en "Estructuras de Datos" (entre otras asignaturas), la cual presenta un alto índice de reprobación y deserción de las actuales carreras de Ingeniería en Computación e Ingeniería Informática del Centro Universitario de Ciencias Exactas e Ingeniería de la Universidad de Guadalajara. (Castellanos et al., 2018), (Beas, 2012).

En toda institución educativa del ámbito público, existe la necesidad constante de ampliar la oferta educativa e incrementar los cupos de admisión. El que una cantidad importante de alumnos ingrese a una carrera y vuelva a cursar varias materias ocasiona un problema al requerir más profesores que atiendan la demanda, en lugar de orientar esos profesores a impartir clase a una mayor cantidad de alumnos que pudieran ingresar.

Se requiere disminuir el índice de reprobación y deserción, detectando el grado de madurez de los alumnos y así elevar el nivel de aprovechamiento.

Objetivos

El objetivo general de esta investigación es medir los antecedentes adquiridos en el bachillerato previo a cursar la unidad de aprendizaje de Programación para determinar el nivel de gestión de conocimiento que posee y que incide en el aprendizaje de este curso.

Los objetivos específicos de este proyecto son los siguientes:

- Aplicar a los alumnos de los ciclos 2020A y 2020B un instrumento para conocer los antecedentes académicos de quienes cursan esta materia
- Analizar e interpretar los datos recabados mediante una aplicación web con base de datos desarrollada para este proyecto, para obtener una muestra representativa.
- Determinar los factores que inciden en el grado de madurez académico de los alumnos, y por lo tanto en el nivel de aprovechamiento de dicha unidad de aprendizaje.

Marco teórico

"El conocimiento, tal como se le concibe hoy, es el proceso progresivo y gradual desarrollado por el hombre para aprehender su mundo y realizarse como individuo, y especie." (Sáez, 1991).

"La gestión del conocimiento es un proceso lógico, organizado y sistemático para producir, transferir y aplicar en situaciones concretas una combinación armónica de saberes, "experiencias, valores, información contextual y apreciaciones expertas que proporcionan un marco para su evaluación e incorporación de nuevas experiencias e información." (Davenport y Prusak, 2001).

Modelo General Knowledge Management Maturity Model (GKMMM)

Es un modelo descriptivo, que expone los atributos esenciales que caracterizan una organización en un nivel particular de la madurez de la gestión del conocimiento. También es un modelo normativo en el que

las prácticas principales caracterizan los tipos ideales de ambiente que deberían esperarse en una organización que esté implementando gestión del conocimiento. El Modelo General de Madurez de GKMMM propuesto adaptó cinco niveles de madurez de CMM (Capability Maturity Model) llamados, de menor a mayor nivel de madurez: inicial, conciencia, definido, gestionado y optimizado (Pee et al., 2006).

Sistemas de apoyo a la toma de decisiones

Una decisión es el resultado de un proceso por el cual se elige entre distintas alternativas. Un Sistema de Apoyo a las Decisiones (DSS por sus siglas en inglés: Decision Support System), es un sistema informático que sustenta el proceso de toma de decisiones, lo cual implica la utilización de datos y modelos para la generación, la estimación, la evaluación y/o la comparación sistemática de alternativas, ayudando a los responsables de la toma de decisión a reunir inteligencia, generar opciones y tomar decisiones. Los DSS constituyen una herramienta para tomar decisiones válidas, con diferentes enfoques y en una amplia gama de actividades. Pueden emplearse para obtener información que revele los elementos destacados de los problemas y las relaciones entre ellos, así como para identificar, crear y comunicar cursos de acción disponibles y alternativas de decisión (Torres et al., 2014).

Algoritmo de Aprendizaje Automático K-NN

En los últimos años, en la rama del aprendizaje automático, han aparecido problemas de la vida práctica donde se hace necesario modelar la predicción a través de estructuras complejas en su salida. Estos problemas están relacionados con la necesidad de predecir valores de salidas en forma de grafos, jerarquías, vectores con valores reales o binarios entre otras. Como término general este tipo de tarea es conocida como problemas de predicción estructurada (Nowozin et al., 2014).

El algoritmo de aprendizaje automático K-Vecinos más Cercanos ("K-Nearest Neighbors", K-NN) está considerado uno de los de mejor desempeño en el ámbito del aprendizaje automático, aunque con alto costo computacional. Al mismo tiempo, K-NN construye un modelo sencillo para resolver problemas de predicción, el cual está basado en los objetos del conjunto de datos de entrenamiento. En K-NN para un nuevo ejemplo, se asigna al atributo de la clase, el valor medio de las clases de los objetos que se encuentran dentro de la frontera de los vecinos más cercanos. En general la distancia Euclidiana es empleada para medir la proximidad entre el nuevo ejemplo y los que conforman el conjunto de datos de entrenamiento (Cover, 1967).

La Universidad de Guadalajara y su modelo educativo.

Actualmente, esta institución educativa (www.udg.mx) se desarrolla a partir de un modelo de Red Universitaria que se integra por varios centros universitarios y escuelas de nivel bachillerato. La calidad de sus programas educativos cuenta con el reconocimiento de organismos nacionales e internacionales (UdeG, 2021).

CUCEI (www.cucei.udg.mx) es uno de los centros universitarios que forman la red, atendiendo a la fecha a más de 15,000 alumnos mediante 18 licenciaturas y 20 posgrados. A través de un modelo organizativo de carácter departamental, desarrolla y genera conocimiento como elemento central para el cumplimiento de sus funciones de docencia, investigación y extensión universitaria. Agrupa y dirige 12 departamentos a través de las divisiones de Ciencias Básicas, Ingenierías y de Electrónica y Computación (CUCEI, 2021).

La presente investigación se hace con estudiantes de algunas de las licenciaturas que se ofertan en este plantel educativo y que cursan la asignatura de Programación.

Modelo Educativo del Sistema de Educación Media Superior de la Universidad de Guadalajara

La Universidad a través del Sistema de Educación Media Superior (SEMS), cubre el mayor espacio educativo de Enseñanza Media Superior en el Estado de Jalisco, ofreciendo una formación integral y de calidad a bachilleres y técnicos. Su oferta académica se divide en varios modelos curriculares, que son los siguientes:

1. Bachillerato general. En modalidad escolarizada o semiescolarizada. "La propuesta de este modelo y que se estudia bajo el modelo de competencias, gira en torno a 5 ejes curriculares: comunicación, pensamiento matemático, comprensión del ser humano y ciudadanía, comprensión de la naturaleza y formación para el bienestar (SEMS, 2021).
2. Bachillerato Tecnológico. Con quince opciones para diversas áreas del conocimiento, incluyendo entre ellas, el Bachillerato Tecnológico en Desarrollo de Software que se imparte en uno de los planteles.
3. Carrera Tecnológica. Con trece opciones para diversas áreas del conocimiento, incluyendo entre ellas, al Tecnólogo Profesional en Informática y al Tecnólogo Profesional en Sistemas Informáticos, mismas que se imparten en la Escuela Politécnica de la Universidad de Guadalajara.

Todas las anteriores tienen un nivel de escolaridad reconocido como a nivel de bachillerato y posteriormente sus egresados pueden estudiar una licenciatura, tales como las que cursan los estudiantes que se tomaron como parte de la muestra de esta investigación.

Sin embargo, la Universidad de Guadalajara admite aspirantes de diversas instituciones educativas de la región y de diversas partes del país. En este estudio se hace un análisis de la información para descubrir si es un factor determinante el que los alumnos que llegan a estudiar una licenciatura en esta misma universidad, hayan estudiado alguna de las modalidades de bachillerato enunciadas anteriormente.

Metodología

El estudio se enmarca en una investigación descriptiva con la cual se pretende medir o recoger información de manera conjunta o independiente sobre los aspectos a evaluar, tendrá un enfoque cuantitativo (cuestionario), el alcance final será de tipo comparado, ya que el propósito fundamental es determinar el nivel de madurez de los alumnos provenientes de diferentes instituciones.

Fases del proyecto

Para realizar la investigación, se llevaron a cabo las siguientes fases:

1. Identificar los criterios que deben de cumplir los encuestados, aplicando el cuestionario a los alumnos que cursan la unidad de aprendizaje de Programación
2. Motivar a los alumnos para que participen en el experimento.
3. Implementar el "Sistema de Detección de Madurez para Aprender Programación", como una aplicación web con base de datos y utilizando componentes de software libre en lenguaje Java.
4. Almacenar en el sistema la información necesaria para este proyecto.
5. Realizar un sondeo sobre los datos recabados de la totalidad de los estudiantes siendo seleccionados de acuerdo a las respuestas obtenidas.
6. Analizar e interpretar los datos recabados en el sistema, para posteriormente definir cómo efectuar un análisis automático de la información.
7. Generar mediante el sistema reportes gráficos y tabulares, con datos de utilidad para la toma de decisiones.
8. Determinar mediante el sistema el grado de madurez curricular de los alumnos, identificando en qué nivel se encuentran acorde al modelo G-KMMM.

La Encuesta

La medición de la madurez de la Gestión del conocimiento se realiza a través de instrumentos tipo encuesta que se aplican a los alumnos de la unidad de aprendizaje de Programación.

Se realiza una investigación con un método cuantitativo estadístico, mediante la técnica del cuestionario y la interpretación de los datos recabados.

El paradigma positivista (cuantitativo) busca descubrir el conocimiento a partir de relaciones causa-efecto con las que pretende controlar, explicar y predecir hechos. El investigador busca la neutralidad y hace que prevalezca la objetividad. Este instrumento se centra en aspectos observables que sean posibles de cuantificar y sean libres de valores (Barbera, 2008). Es importante aclarar, que este método se aplicó a la mayoría de las preguntas y para algunas otras se recabó información con el objetivo de conocer algunas de las opiniones de los estudiantes. El detalle se puede apreciar en la **Tabla 1**.

Las preguntas abiertas son más fáciles de formular; sin embargo, la dificultad aparece a la hora de poder tratar la información recabada, ya que se deben reducir las contestaciones desiguales a unas categorías significativas cuantificables.

Para Presser y Schuman (1981) las preguntas abiertas son esenciales para conocer el contexto del sujeto que contesta el cuestionario y para redactar después las alternativas a ofrecer en las preguntas categorizadas, especialmente cuando no se pueden presumir con antelación, las posibles opiniones y reacciones de la población a que se va a aplicar el cuestionario

La investigación científica, desde el punto de vista cuantitativo es un proceso sistemático y ordenado que se lleva a cabo siguiendo determinados pasos. Planear una investigación consiste en proyectar el trabajo de acuerdo con una estructura lógica de decisiones y con una estrategia que oriente la obtención de respuestas adecuadas a los problemas de indagación propuestos. Pese a tratarse de un proceso metódico y sistemático, no existe un esquema completo, de validez universal, aplicable mecánicamente a todo tipo de investigación. No obstante, sí es posible identificar una serie de elementos comunes, lógicamente estructurados, que proporcionan dirección y guía en el momento de realizar una investigación, los cuales se pueden organizar en fases y etapas. Es menester aclarar entonces que los pasos que se señalan no constituyen una guía inflexible, puesto que es posible que, en cada investigación en particular, algunos de ellos se superpongan, otros sean intercambiables, no sigan la secuencia lineal preestablecida o simplemente en ciertos casos resulten innecesarios. Las etapas que se desarrollan son las siguientes: fase conceptual, fase de planeación y diseño, fase empírica, fase analítica y fase de difusión (Monje, 2017).

La encuesta consta de 26 preguntas, en la cual la mayoría de ellas deben responderse con un SÍ o un NO, y algunas con respuestas abiertas. Los alumnos participantes se seleccionaron al azar, siendo alumnos que cursan materias ofertadas en la División de Electrónica y Computación (DIVEC).

El cuestionario elaborado para este propósito fue sometido a una validación por parte del Cuerpo Académico de Investigación Educativa en Tecnologías de la Información, con siglas CA998 de la Universidad de Guadalajara; la totalidad de los integrantes del grupo y dos profesores colaboradores del Departamento de Ciencias Computacionales de la DIVEC participaron en la revisión del mismo.

Entre los aspectos que se consideraron para aprobar este instrumento están algunos, tales como el empleo de términos comprensibles para los participantes, la confiabilidad de los resultados o la correcta secuenciación de los reactivos entre otros.

El instrumento utilizado se aplicó a una población de alumnos entre los 17 y los 27 años, en una encuesta que se llevó a la práctica al término de los ciclos escolares 2020A y 2020B a un total de 144 estudiantes de las carreras de las Ingenierías en Computación, en Informática, en Biomédica, en Comunicaciones y Electrónica y en Civil, obteniendo de manera anónima información sobre su experiencia educativa previa en el bachillerato y sobre sus antecedentes en el área de la programación.

La muestra se consideró adecuada y suficiente por ser una cantidad que ya es representativa y sin ser tan grande para que se dificulte su manejo, con una participación voluntaria de estudiantes que mostraron su interés en aportar su opinión, que resulta benéfica no solo para alumnos de un determinado género, edad, condición social, licenciatura o algún otro factor específico de manera que pueda ser aplicada a jóvenes de nuevas generaciones.

La encuesta se diseñó con un formulario de Google. Las preguntas que se incluyen en la encuesta se encuentran referidas en la **Tabla 1**.

Pregunta	Opciones	Tipo de respuesta	Método
1. Selecciona la SECCIÓN a la que corresponda tu grupo	Lista desplegable: 1. D12, I5882, ARANGUREN NAVARRO ITZEL NIASANDIU, MJ 1100-1255 2. D34, I5882, ARANGUREN NAVARRO ITZEL NIASANDIU, MJ 0700-0855 ... 41. D22, I5882, ZAMORA RAMOS VICTOR MANUEL, LI 1100-1255	Obligatoria	Cuantitativo
2. Anota tu código de estudiante	Respuesta corta	Obligatoria	Cualitativo
3. Selecciona la clave de tu Carrera	Lista desplegable: 1. INCO 2. INNI 3. INCE	Obligatoria	Cuantitativo

	4. INBI 5. ICIV		
4. Anota el nombre de tu escuela de procedencia de nivel bachillerato (sin usar abreviaturas)	Respuesta corta	Obligatoria	Cualitativo
5. Anota el nombre del estado donde se ubica tu escuela de procedencia (sin usar abreviaturas)	Varias opciones: Jalisco Otra	Obligatoria	Cuantitativo
6. Anota el nombre del municipio donde se ubica tu escuela de procedencia (sin usar abreviaturas)	Respuesta corta	Obligatoria	Cualitativo
7.1 Anota el año de inicio de la preparatoria (p. ejem "2015")	Respuesta corta	Obligatoria	Cuantitativo
7.2 Elige el mes de inicio de la preparatoria	Lista desplegable: 1. Enero 2. Febrero ... 12. Diciembre	No obligatoria	Cuantitativo
7.3 Anota el año de fin de la preparatoria	Respuesta corta	Obligatoria	Cuantitativo
7.4 Elige el mes de fin de la preparatoria	Lista desplegable: 1. Enero 2. Febrero ... 12. Diciembre	No obligatoria	Cuantitativo
8. Duración oficial del bachillerato (en años)	Lista desplegable: 1.1 2. 2 3. 3 4. 4 5. 5	Obligatoria	Cuantitativo
9. ¿Tu escuela de bachillerato es pública o privada?	Varias opciones Pública Privada	Obligatoria	Cuantitativo
10. ¿Tu bachillerato pertenece al sistema de la UdeG o incorporada a la UdeG?	Varias opciones: Si No	Obligatoria	Cuantitativo
11. ¿Tu bachillerato es técnico (CONALEP, CBETIS, CETIS, CECYTEJ, etc)?	Varias opciones: Si No	Obligatoria	Cuantitativo
12. Ciclo de admisión de la carrera actual	Varias opciones: 2018B 2019A 2019B 2020A 2020B Otra	Obligatoria	Cuantitativo
13. ¿Es la primera vez que tomas esta materia?	Varias opciones: Si No	Obligatoria	Cuantitativo
14. ¿Habías cursado materias de programación antes de tomar este curso por primera vez?	Varias opciones: Si No	Obligatoria	Cuantitativo
15. Enlista los Lenguajes de programación que ya conocías (por tu cuenta o como parte de un curso de tu bachillerato) antes de tomar por primera vez este curso; en caso de necesitar citar "otros", utiliza comas para señalar más de un lenguaje	Casillas: C C++ C# Java Python Otra	Obligatoria	Cuantitativo
16. En base a tu kardex del bachillerato, enlista los nombres de las materias cuyo contenido sea de matemáticas (utiliza "," para separar los nombres de las materias que cites)	Párrafo	Obligatoria	Cuantitativo
17. En base a tu kardex del bachillerato, enlista los nombres de las materias cuyo contenido sea de programación y no sólo paquetería de Office (utiliza "," para separar los nombres de las materias que cites); en caso de no haber cursado ninguna de programación, escribe "Ninguna"	Párrafo	Obligatoria	Cuantitativo
18. ¿Has participado en concursos u olimpiadas de programación?	Varias opciones: Si No	No Obligatoria	Cuantitativo
19. ¿Has participado en concursos u olimpiadas de matemáticas?	Varias opciones: Si No	No Obligatoria	Cuantitativo
20. ¿Qué edad tienes al momento de responder esta encuesta?	Varias opciones: 18 19 20 21 22 23 Otra	No Obligatoria	Cuantitativo
21. ¿Eres autodidacta?	Varias opciones: Si No	Obligatoria	Cuantitativo
22. Cita el nombre de algún libro que hayas leído recientemente (en los últimos 3 meses) independientemente del género que haya sido, pero que no haya sido	Párrafo	Obligatoria	Cuantitativo

necesario leerlo para resolver una tarea escolar (puedes citar incluso un libro necesario para tu carrera, solo si lo leíste por el gusto de aprender más temas, que lo que necesitaste para tus tareas); en caso de no haber leído ninguno recientemente, escribe "Ninguno"			
23. ¿De cuántas horas a la semana dispones para la realización de tareas escolares? Escribe un número del 0 al 99	Respuesta corta Número: 0 a 99. En caso contrario está fuera del rango	Obligatoria	Cuantitativo
24. ¿Cuántas materias inscribiste en este ciclo escolar? Escribe un número del 1 al 12	Respuesta corta Número: 1 a 12. En caso contrario está fuera del rango	Obligatoria	Cuantitativo
25. Comenta lo que consideres que es una ventaja (o ventajas) que tienes para aprender esta materia, y que, posiblemente, otros compañeros tuyos no poseen. En caso de no haberlas escribe "Ninguna"	Párrafo	Obligatoria	Cuantitativo
26. Comenta lo que consideres que es una desventaja (o desventajas) que tienes para aprender esta materia. En caso de no haberlos escribe "Ninguna"	Párrafo	Obligatoria	Cuantitativo

Tabla 1. Contenido de la encuesta. Fuente: Diseño propio.

En las **Figuras 1 y 2** se observa parte del formulario que se aplicó para elaborar la presente investigación al término de los ciclos escolares, cuando ya había concluido el periodo oficial de clases.



Figura 1. Encuesta para recabar información para el estudio.

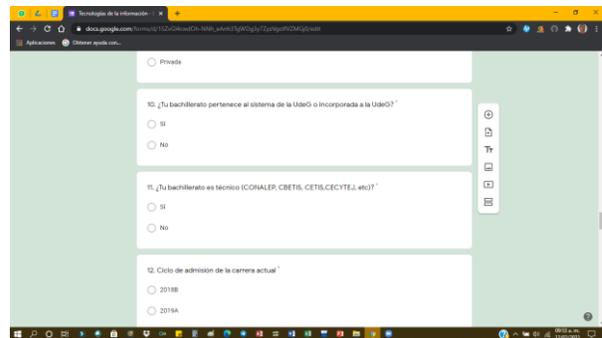


Figura 2. Ejemplo de las preguntas números 10 y 11 de la encuesta.

A partir de la información recabada en el formulario de Google que fue utilizado para este fin, se obtiene un archivo en formato de Excel, generado como insumo para la aplicación web diseñada e implementada a la medida para este proyecto.

Sistema de Detección de Madurez para Aprender Programación (SDMAP)

Este sistema fue implementado con herramientas gratuitas como Java, el acceso a los datos se realiza mediante un gestor de MySQL, además que se utilizó el sistema distribuido ApacheTomcat. Se instaló en un servidor y funciona de manera concurrente, ya que está diseñado para recibir peticiones simultáneas de cualquier cantidad de clientes delegando a un nuevo hilo dinámico el procesamiento de cada conjunto de registros que reciba.

La aplicación web permite:

- Recolectar la información de estudiantes obtenida mediante encuesta.
- Capturar las calificaciones finales del curso de Programación obtenidas por los estudiantes.
- Generar reportes con datos de utilidad para la toma de decisiones.
- Interpretar los datos de forma gráfica.
- Determinar la madurez que un estudiante tuvo en particular, previo a incursionar en la materia de Programación.

El SDMAP cuenta con los siguientes módulos:

- Administración y autenticación de usuarios: estudiantes y profesores.
- Interpretación de datos:
 - Representación de resultados con:
 - Diagramas de dispersión
 - Histogramas
 - Determinación de nivel de madurez:
 - Algoritmo de aprendizaje automático (KNN)

El módulo de determinación de nivel de madurez se implementó con la técnica KNN en la que un nuevo caso se clasificaría en la clase más frecuente a la que pertenezcan sus "K" vecinos más cercanos. El algoritmo, por tanto, es intuitivo, lo que unido a su fácil implementación hace que sea un paradigma clasificatorio muy extendido. Este algoritmo se eligió debido a que permite clasificar registros nuevos basándose en el cálculo de la distancia euclidiana en un mapa cartesiano N-dimensional, y en este proyecto se trabaja con 5 dimensiones, una para cada pregunta que se determinó como relevante, y se detalla en los resultados estas mismas.

Resultados

Análisis y validación de datos

Durante la implementación del sistema, antes de que se pusiera a punto, se analizaron los datos manualmente mediante hojas de cálculo de Excel, con el propósito de encontrar tendencias entre las respuestas de la encuesta y las calificaciones finales en el curso de Programación.

Para encontrar dichas tendencias, primero se clasificaron los registros por cada tipo de pregunta de la encuesta y para cada una de ellas se determinaron dos conjuntos de estudiantes:

- con fortalezas (o ventajas) que inciden en una mejor calificación final;
- con debilidades (o desventajas).

A continuación, para cada conjunto de estudiantes se promediaron sus calificaciones obtenidas del curso de Programación y se calcularon 3 tipos de promedios:

- de calificaciones por cada profesor;
- de calificaciones entre todos los profesores; y
- general por ciclo escolar.

En la **Tabla 2** se ilustra un análisis de fortalezas y debilidades de la pregunta número 14, donde para los fines de este ejemplo, cada profesor se encarga de impartir clase a un grupo de alumnos:

14. ¿Habías cursado materias de programación antes de tomar este curso por primera vez?				
Grupo	Sí	No	Diferencia en calificación del ciclo escolar	Porcentaje de incidencia del ciclo escolar
1	76.75	75.12	1.63	2.169861555
2	98.92	98.27	0.65	0.6614429633
3	76.67	78.83	-2.16	-2.740073576
4	87	79.47	7.53	9.475273688
5	94.86	ninguno	no aplica	no aplica
6	91.17	76.22	14.95	19.61427447
Promedio de grupos	87.56	81.58	5.98	7.329639708
Promedio general	91.53	80.97	10.56	13.04186736
Conclusión "Promedio de grupos"	Obtienen un promedio superior quienes cursaron materias de programación antes de ingresar a la unidad de aprendizaje de Programación por primera vez, que quienes llegan con ninguna experiencia, excepto en el grupo 3, ya que hay una diferencia negativa en calificación entre los conjuntos "Sí" (fortalezas) y "No" (debilidades)			
Conclusión "Promedio general"	En el "promedio general" se confirma la tendencia observada en el "promedio de grupos"			

Tabla 2. Ejemplo de la pregunta número 14 de la encuesta en el ciclo 2020A. Fuente: Diseño Propio.

Análogamente al análisis descrito anteriormente, se estudiaron las demás preguntas de la encuesta tanto para el ciclo escolar 2020A como para el 2020B, obteniendo los resultados siguientes:

- Para algunas preguntas la tendencia se confirmó en ambos ciclos escolares;
- Para otras preguntas la tendencia fue contradictoria en el segundo ciclo escolar;
- Se buscó la tendencia generando promedios en independencia del ciclo escolar;
- Se descartaron las preguntas donde la tendencia era contradictoria entre ambos ciclos escolares;
- Del total de las preguntas sólo en 5 se detectaron resultados relevantes.

Las preguntas donde se observó una mayor diferencia en calificaciones finales obtenidas por los estudiantes tienen que ver (de mayor a menor), con los siguientes:

- Identificar el perfil del encuestado enfocado en su carrera profesional, es decir Computación e Informática, en contraste con las demás carreras, esto es, la pregunta número 3;
- El ciclo de admisión, esto es, la pregunta número 12;
- Los conocimientos previos en programación habiendo cursado materias relacionadas en su bachillerato, esto es, la pregunta número 14;

- La duración de la currícula de bachillerato, esto es, la pregunta número 8; y
- El dominio de más lenguajes de programación, esto es, la pregunta número 15.

La siguiente información se obtiene conjuntando los resultados de ambos ciclos escolares, quedando como se ilustra en la **Tabla 3**. Las cinco preguntas representan las fortalezas que mostraron los estudiantes, para obtener mejores resultados en la materia de Programación, y que son la base para determinar el grado de madurez.

Número de Pregunta	Diferencia en calificación de los ciclos 20A y 20B	Porcentaje de incidencia de los ciclos 20A y 20B
3	13.48861111	34.09778739
12	10.25333333	25.91934612
14	7.861944444	19.87416702
8	4.048333333	10.23376003
15	3.906388889	9.874939436
Total	39.55861111	100

Tabla 3. Resultado de la integración de los dos ciclos escolares. Fuente: Diseño propio.

Resultados del SDMAP

Los resultados obtenidos del análisis de las preguntas de la encuesta se contrastaron y se verificaron contra los que presentó el sistema ya implementado. Por ejemplo, en la **Figura 3** se representa un reporte con los porcentajes de diferencia por grupo para la pregunta número 14. Para cada una de las preguntas relevantes el sistema genera un reporte con el mismo tipo de información.

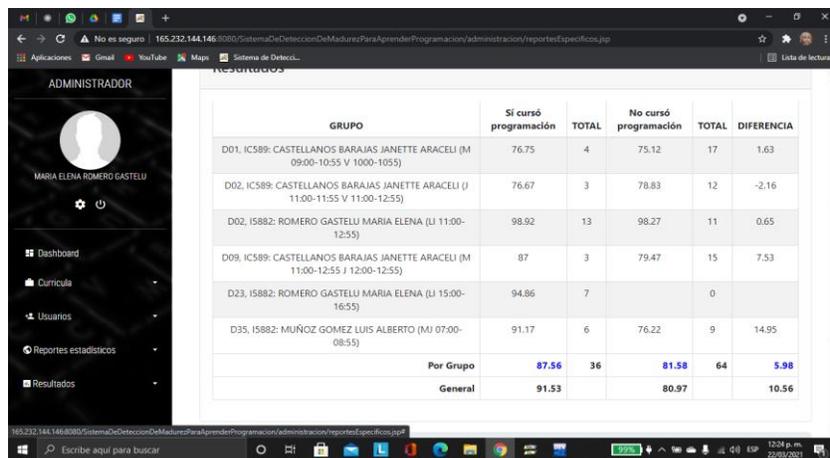


Figura 3. Resultados de la pregunta número 14, del ciclo escolar 2020A.

Además de estos reportes, el sistema para cada una de las preguntas permite seleccionar el tipo de gráfico que sea más conveniente, ya sea de dispersión (**Figura 4**) o de histograma (**Figura 5**).

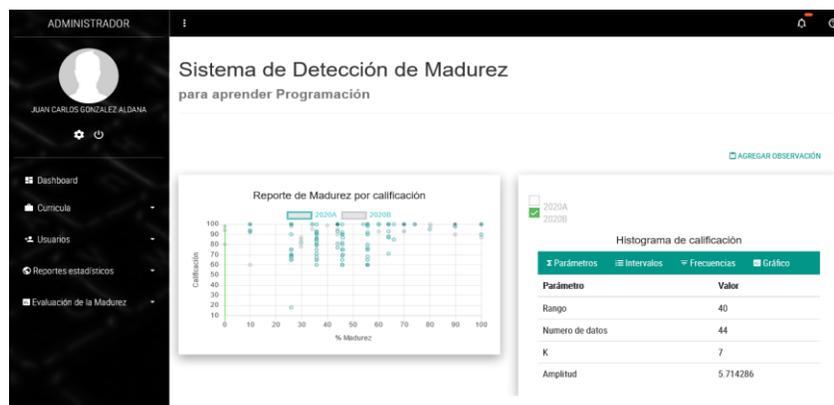


Figura 4. Análisis de dispersión y datos del histograma de calificaciones

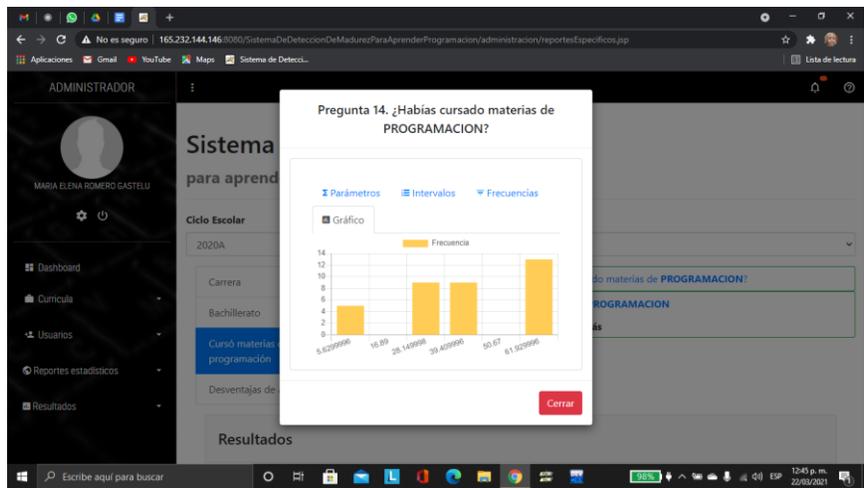


Figura 5. Reporte estadístico de tipo histograma.

Por otra parte, el sistema cuenta con el módulo para determinar el grado de madurez, que hace el cálculo a partir de la captura de los datos de las cinco preguntas consideradas relevantes y que son analizados por el sistema con el algoritmo KNN, mismo que evalúa el nivel de madurez que corresponde de acuerdo a los criterios determinados. En la **Figura 6** se ilustra el formulario para capturar los datos.

Field Label	Value
Ciclo Escolar Actual	2020A
Pregunta 3. Selección de la Carrera	INNI - ING INFORMATICA
Pregunta 8. Duración oficial del bachillerato (en años)	3
Pregunta 12. Ciclo de admisión de la carrera actual	2019B
Pregunta 14. ¿Habías cursado materias de PROGRAMACION?	SI
Pregunta 15. Cantidad de lenguajes de PROGRAMACION que conoce	4

Figura 6. Interfaz para la evaluación de un registro no almacenado.

En la **Figura 7** se muestra el resultado de la evaluación de madurez, tras aplicar el algoritmo K-NN.

Evaluación Resultante

[96.75, 87.24615478515625, 95.41071319580078, 93.2203369140625, 89.70755004882812]

Clasificado como Nivel 'OPTIMIZADO' [87,100]

Figura 7. Evaluación de madurez con K-NN.

A continuación se describe el procedimiento con el que se determinaron los niveles de madurez para esta investigación:

- Se realizó una limpieza de datos, de los 144 registros originales con que contaba la encuesta, quedando un total de 140 de los mismos, todos con patrones comunes que permitieron clasificarlos.
- Con el análisis de la información ya almacenada se realizó la creación de un conjunto de datos descriptivos por estudiante (Dataset en **Tabla 4**). Éste sirvió para la alimentación del algoritmo de aprendizaje automático (KNN) capaz de predecir una etiqueta de clasificación para cada estudiante.
- Los datos de los 140 registros del Dataset, quedaron clasificados con colores y con números. Los colores representan las fortalezas y los valores numéricos los promedios de calificaciones de aquellos estudiantes que cuentan con alguna fortaleza como se observa en la **Tabla 4**. Por otro lado, en la **Tabla 5** se describen las fortalezas de cada una de las preguntas relevantes.

Cantidad Fortalezas	Pregunta 3	Pregunta 8	Pregunta 12	Pregunta 14	Pregunta 15	Calificación Final	Clase K-NN (etiqueta)
5	96.75	94.07	93.45	93.22	92.07	100	Optimizado

4	96.75	87.25	93.45	93.22	92.07	90	Optimizado
3	96.75	87.25	93.45	93.22	87.08	100	Gestionado

1	96.75	87.25	85.47	84.22	87.08	81	Gestionado
2	81.6	94.07	85.47	93.22	87.08	99	Definido

2	81.6	94.07	93.45	84.22	87.08	70	Definido
2	81.6	87.25	93.45	93.22	87.08	98	Conciencia

1	81.6	87.25	85.47	93.22	87.08	78	Conciencia
1	81.6	87.25	93.45	84.22	87.08	100	Inicial

0	81.6	87.25	85.47	84.22	87.08	98	Inicial

Tabla 4. Muestra parcial del Dataset que la técnica K-NN necesita para predecir la clase a la que corresponde cada uno de los nuevos registros que fuesen capturados.

No. pregunta relevante	Descripción de la pregunta	El estudiante cuenta con fortaleza si...
3	Perfil de carrera del estudiante	Está inscrito a las carreras de Informática y Computación
8	Duración del bachillerato	La duración es de 4 o más años curriculares
12	Ciclo de admisión a la carrera que cursa el estudiante	Que sea anterior al ciclo escolar en que ingresa a la unidad de aprendizaje de Programación
14	Haber ejercido la programación durante su bachillerato	Cursó alguna asignatura relacionada a la Programación
15	Cantidad de lenguajes de programación	Conoce 2 o más lenguajes de programación

Tabla 5. Tabla de Fortalezas por estudiante por pregunta relevante. Fuente: Diseño propio.

- Las etiquetas que aparecen en el dataset de la **Tabla 4** se tomaron como base del modelo G-KMMM sobre la gestión del conocimiento (**Tabla 6**).

Nivel de Madurez G-KMMM	Pregunta 3	Pregunta 8	Pregunta 12	Pregunta 14	Pregunta 15	Rango de calificación predecible
Inicial	N	N	S/N	N	N	18 a 100
Conciencia	N	N	S/N	S/N	N	60 a 100
Definido	N	S/N	S/N	S/N	S/N	68 a 100
Gestionado	S	N	S/N	S/N	N	81 a 100
Optimizado	S	S/N	S/N	S/N	S	87 a 100

Tabla 6. Niveles de madurez G-KMMM. Fuente: Diseño propio.

- Una vez interpretados los resultados del dataset y resumidos como se muestra en la **Tabla 6** y tomando como referencia la descripción de cada una de las fortalezas de la **Tabla 5**, se determinó la descripción de cada uno de los niveles de madurez (**Tabla 7**).

Nivel de madurez	Descripción
Inicial	El estudiante cuenta con poco o ningún conocimiento previo que pudiera aplicar; invariablemente no está inscrito a las carreras de las TICs, la duración de su bachillerato es de 3 años o menos, no cursó asignatura alguna relacionada a la Programación y si acaso conoce un y solo un lenguaje de programación.
Conciencia	El estudiante cuenta con algún conocimiento previo que le es de ayuda, como cursar más asignaturas en licenciatura antes de Programación, o bien, alguna asignatura relacionada específicamente a la Programación; invariablemente no está inscrito a las carreras de las TICs, la duración de su bachillerato es de 3 años o menos y si acaso conoce un y solo un lenguaje de programación.
Definido	El estudiante cuenta con dos a cuatro fortalezas previas que le son de ayuda, como lo son, la duración de su bachillerato de 4 años o más y/o, su ciclo de admisión al haber cursado asignaturas de licenciatura antes de Programación y/o, haber cursado alguna asignatura relacionada a la Programación y/o conocer 2 o más lenguajes de Programación; invariablemente no está inscrito a las carreras de las TICs. Lo define en sí su mayor experiencia académica, por lo que es poco probable el no acreditar la unidad de aprendizaje, acorde a su rango de calificación predecible.
Gestionado	El estudiante está inscrito a las carreras de las TICs, lo cual incide en su interés por adquirir los conocimientos de la unidad de aprendizaje; invariablemente la duración de su bachillerato es de 3 años o menos y si acaso conoce un y solo un lenguaje de programación, pero puede contar con otras fortalezas como lo son su ciclo de admisión y/o haber cursado alguna asignatura relacionada a la Programación.
Optimizado	El estudiante está inscrito a las carreras de las TICs, lo cual incide en su interés por adquirir los conocimientos de la unidad de aprendizaje y además conoce 2 o más lenguajes de programación, lo cual incide en el mejor rango de calificación predecible; además puede contar con otras fortalezas como la mayor duración de su bachillerato y/o su ciclo de admisión y/o haber cursado alguna asignatura relacionada a la Programación.

Tabla 7. Descripción de niveles de madurez. Fuente: Diseño propio.

Conclusiones y trabajo futuro

Los datos recabados, analizados e interpretados de diferentes formas con el sistema de tipo DSS para la toma de decisiones “Sistema de Detección de Madurez para Aprender Programación” arrojó resultados de valor, ya que permitió identificar fortalezas, así como las preguntas relevantes en este estudio, con las cuales se pueden realizar otros estudios posteriores. Con este sistema se obtuvieron reportes, diagramas de dispersión e histogramas, así como la determinación de los niveles de madurez para estudiantes que cursan o cursarán la materia de Programación.

Se logró determinar el grado de madurez de los 140 estudiantes y con los patrones comunes encontrados con los datos de la encuesta se construyó la descripción para cada uno de estos cinco niveles basados en el modelo GKMMM (**Tabla 7**) y que son coherentes con las fortalezas identificadas en las preguntas relevantes.

Con un mayor volumen de datos que se obtengan de los siguientes ciclos escolares se puede enriquecer el algoritmo KNN y se podrán tomar decisiones a futuro para mejorar la investigación. Podemos decir que el modelo GKMMM como el KNN fueron útiles para este propósito. Es un trabajo relevante que no cuenta con antecedentes a la fecha.

A continuación, se enlistan las conclusiones y reflexiones por cada una de las cinco preguntas más relevantes de esta investigación:

Pregunta número 3 “Selecciona la clave de tu Carrera”: Se comprueba que los estudiantes con un perfil afín a las TIC 's, alumnos de las carreras de Informática y Computación, tienen una ventaja sobre los estudiantes que no la tienen.

Pregunta número 12 “Ciclo de admisión de la carrera actual”: Se comprueba que las personas que son de ciclos anteriores al ciclo escolar en que se aplica la encuesta, tienen más fortalezas y una mejor formación previa, que las nuevas generaciones de estudiantes. Aquellos con fortalezas son estudiantes que tienen rangos de edades mayores; por otro lado, aquellos con debilidades coinciden por su edad, con haber sido involucrados en una reforma educativa cuando cursaron el nivel primaria y secundaria (INEE, 2015).

Pregunta número 14 “¿Habías cursado materias de programación antes de tomar este curso por primera vez?” y la pregunta número 15 “Enlista los Lenguajes de programación que ya conocías (por tu cuenta o como parte de un curso de tu bachillerato) antes de tomar por primera vez este curso; en caso de necesitar citar “otros”, utiliza comas para señalar más de un lenguaje”: Se comprueba que los conocimientos previos en programación influyeron de manera positiva en el aprovechamiento de los estudiantes; algunos

adquirieron conocimientos cursando materias relacionadas a ello en su bachillerato; otros, en el caso de los estudiantes de Ingeniería en Computación que cursan en un semestre previo una materia introductoria a la computación, su experiencia les fortalece. Esto se comprueba con la teoría del diseño de algoritmos computacionales, que dice que mientras más se practica la solución de problemas, es mayor la habilidad de resolver los mismos, puesto que se fortalece el pensamiento computacional (BASOGAIN et al., 2015).

Pregunta número 8 “Duración oficial del bachillerato (en años)”: Se comprueba que mientras más años dura el bachillerato, obtienen mejores resultados los estudiantes en la materia de Programación, posiblemente debido a que cursaron una cantidad mayor de materias. Cabe resaltar que los alumnos provienen de bachilleratos de diferentes sistemas educativos y que cada uno de estos tienen diferentes planes de estudio.

Al ampliar esta investigación se espera que los resultados en el aprovechamiento de la unidad de aprendizaje de Programación se pueden ver mejorados y el índice de deserción tendería a disminuir. Si se logra mejorar las competencias de los estudiantes de las carreras de Ingeniería en Computación e Ingeniería Informática, en lo que respecta al curso de Programación (y otros posteriores), se elevarían los índices de permanencia y egreso, y se obtendrían mejores resultados en procesos de acreditación con evaluaciones realizadas por comités nacionales e internacionales, aumentando con ello el prestigio de la institución.

El presente trabajo se puede complementar con otras investigaciones alternativas, tales como las siguientes:

Debido a que no se encontró una correlación entre los cursos de matemáticas generales del bachillerato y las competencias requeridas para mejorar la práctica de la programación, es necesario, hacer un estudio específico para determinar cuáles son las matemáticas que influyen en los resultados obtenidos de aprendizaje en la asignatura de Programación que se imparte en CUCEI.

Llevar a cabo una investigación en la que se explora cómo el tiempo que el estudiante practica la programación incide en mejores resultados en cursos futuros.

Proponer una solución que fortalezca las competencias de nuestros estudiantes, aplicable antes de que ingresen a cursar dicha unidad de aprendizaje, como por ejemplo: curso propedéutico, asesorías, materias transversales, etc.

Añadir un módulo al sistema informático para que interprete los archivos digitales que contengan la información curricular de cada uno de los estudiantes de su bachillerato de origen, como una boleta de calificaciones, y mediante el algoritmo adecuado se proyecte la calificación final que un estudiante podría obtener en el curso de Programación.

Este trabajo fue posible realizar gracias a la participación de diversos alumnos que aceptaron participar en la encuesta, así como a varios profesores que imparten la unidad de aprendizaje de Programación.

Referencias bibliográficas

- BARBERÁ, E. (2008). Aprender e-learning. Barcelona: Paidós.
- BASOGAIN, Xabier, OLABE, Miguel, OLABE, Juan (2015). Pensamiento Computacional a través de la Programación:
- BEAS, María (2012). Estudio de egresados de la Maestría en Ciencias de la Educación Generación 2007-2009, del Instituto Superior de Investigación y Docencia para el magisterio de la Secretaría de Educación Jalisco. México: Universidad Autónoma de Guadalajara.
- CASTELLANOS, Araceli, ESQUIVEL, Sara, GOMEZ, Abelardo, ORTEGA, Noé y ROMERO, María (2018). Investigación en Competencias Educativas en Educación Superior. Enfoque en las aulas de enseñanza de las tecnologías de la información. Caso Práctico. México: Editorial TRAUCO.
- COVER, T., & HART, P. (1967). Nearest neighbor pattern classification. Information Theory, IEEE Transactions on, 13(1), 21–27.
- CUCEI. Sitio Oficial del Centro Universitario de Ciencias Exactas e Ingenierías. Acerca del CUCEI. En <<http://www.cucei.udg.mx/es/acerca-de>> [Consulta: 20/03/21]
- DAVENPORT, T. y PRUSAK, L. (2001). Conocimiento en acción. Brasil. Prentice Hall.

- Instituto Nacional para la Evaluación de la Educación (2015). Reforma Educativa, Marco normativo. En INEE <https://www.senado.gob.mx/comisiones/educacion/docs/docs_INEE/Reforma_Educativa_Marco_normativo.pdf> [Consulta: 20/03/21]
- LOPEZ, Claudia, TORRENT, Marcela, MOLINA, Florencio, SCHANZ, Sebastián, SANDOVAL, Anahi, SFILIGOY, Juan, SASSAROLI, Marcos (2014). Sistemas de apoyo a la toma de decisiones. Una aplicación en el área de Gestión Universitaria. En SEDICI. <<http://sedici.unlp.edu.ar/handle/10915/42010>> [Consulta: 20/03/21]
- MONJE, Carlos (2017). Metodología de la Investigación Cuantitativa y Cualitativa. Guía Didáctica. En la Universidad Surcolombiana. Facultad de Ciencias Sociales y Humanas <<https://www.uv.mx/rmipe/files/2017/02/Guia-didactica-metodologia-de-la-investigacion.pdf>> [Consulta: 20/03/21]
- NOWOZIN, S., GEHLER, P. V., JANCSARY, J., & LAMPERT, C. H. (2014). Advanced Structured Prediction. MIT Press.
- Paradigma de Aprendizaje. En RED Revista de Educación a Distancia <<https://www.um.es/ead/red/46/Basogain.pdf>> [Consulta: 20/03/21]
- PEE, L., TEAH, H. y KANKANHALLI, A. (2006). Development of a General Knowledge Management Maturity Model. Ponencia presentada en Korean Knowledge Management Society Conference (pp. 17-18). Seúl, Corea del Sur.
- SAEZ, R. El conocimiento científico y el conocimiento mítico asumidos en el método pedagógico. Revista Complutense de Educación. 1991;2:277-95. Disponible en: <http://www.ucm.es/BUCM/revistas/edu/11302496/articulos/RCED9191230277A.PDF>
- SCHUMAN, H. y PRESSER, S. (1981): Questions & answers in attitude survey. Experiments on questions form, wording & context. Academic P., New York. En Sierra Bravo, R., Op. cit., pág. 308.
- SEMS. Sitio Oficial del SEMS. Bachillerato General por Competencias. En <<http://www.sems.udg.mx/bachillerato-general-por-competencias-bgc#:~:text=El%20Bachillerato%20General%20por%20Competencias,y%20orientado%20hacia%20el%20constructivismo>> [Consulta: 20/03/21]
- UdeG. Sitio oficial de la Universidad de Guadalajara. Presentación. En <<http://www.udg.mx/es/nuestra/presentacion>> [Consulta: 20/03/21]