

Aplicación de la inteligencia artificial en la sistematización de procesos educativos. Caso: Sistema de detección de riesgo escolar en ESCOM

Alicia Guillermina Juárez Carrión
Instituto Politécnico Nacional
juca38_alina@hotmail.com

Jorge Cortés Galicia
Instituto Politécnico Nacional
jcortesg@ipn.mx

Ukranio Coronilla Contreras
Instituto Politécnico Nacional
ukraniocc@yahoo.com

Resumen

Para aquellos que son ajenos a las ciencias computacionales desconocen como un sistema de información, tras un procesamiento de datos, puede tomar decisiones y generar resultados asemejándose al proceso del pensamiento humano. En el ámbito de los sistemas computacionales la inteligencia artificial (IA) intenta aproximarse al funcionamiento mental por medio de algoritmos, que son una serie de pasos detallados para resolver un problema y que podrían compararse con el proceso cerebral al realizar actividades cognitivas. Así como las personas captan información a través de los sentidos, los sistemas de información deben ser alimentados con datos, entre los cuales, surgirán las variables para desarrollar los algoritmos.

Las aplicaciones de los sistemas de información basados en la IA, generalmente se usan para la sistematización de procesos que involucran decisiones, por lo tanto el campo de aplicación es vasto. Como un ejemplo de aplicación en el ámbito educativo, se presenta un proyecto en desarrollo, para la opción de titulación curricular en Escuela Superior de

Cómputo del Instituto Politécnico Nacional (ESCOM – IPN), al que se le denomina trabajo terminal, que consiste en un sistema de detección de riesgo escolar de alumnos de nuevo ingreso en la ESCOM, con el fin de proveerles la atención necesaria como medidas de prevención ante dicho riesgo (deserción y/o bajo rendimiento).

Palabras clave: inteligencia artificial, riesgo escolar, sistematización.

Introducción

En la vida cotidiana desempeñamos una serie de actividades que resuelven las necesidades diarias, desde encender la luz al levantarnos, arreglarnos, desayunar, tomar el autobús, etc. Estas actividades son rutinarias y sencillas por lo que olvidamos que existió un proceso de aprendizaje y entrenamiento para realizarlas; en dicho proceso hay implícitos una serie de pasos que van en incremento conforme a la complejidad de la actividad. Cuando el grado de complejidad de la actividad es alto o bien se requiere analizar el resultado de la actividad, el proceso se hace consciente porque requiere de revisar diferentes maneras para realizarlo e incluso en algunos casos, se involucran una serie de subprocesos que representan a lo que comúnmente denominamos: “resolución de un problema”.

En la búsqueda de la solución de un problema necesitamos información del entorno que obtenemos a través de nuestros sentidos. Esta información adquiere significado gracias al proceso fisiológico del sistema nervioso gobernado por el cerebro; por ejemplo cuando la información entra por medio del sentido de la vista, la corteza visual del cerebro capta las figuras provenientes de imágenes y las irradia a otras partes por diferentes vías de procesamiento, como la vía ventral que procesa el color y la forma, el ¿qué? (que es lo que estoy viendo y nombra las cosas) y la vía dorsal, que procesa la profundidad y el movimiento, el ¿dónde? (ubica los objetos en un espacio físico). Asimismo esta información se procesa en el sistema límbico cuya función se relaciona con las respuestas

emocionales, el aprendizaje y la memoria; en este ejemplo resulta la respuesta emocional al ver una imagen.

La combinación de los tres procesos (ventral, dorsal y límbico) hace que la información adquiera significado en diferentes formas; cuando surgen cuestionamientos acerca de lo que estamos viendo el cerebro procesa en paralelo los segmentos de información que contestan al grupo de preguntas y crea un modelo mental (Globalemotion, 2011; Medphys, 2010; Neurocritic, 2012).

El ejemplo anterior sólo muestra a grosso modo y en una mínima parte cómo funciona nuestro cerebro, tema vasto en el que intervienen la neurociencia y disciplinas relacionadas, que nos hace reflexionar acerca de una herramienta similar al cerebro humano, la computadora. En la actualidad muchas de las actividades de la vida cotidiana a nivel personal, laboral o profesional las realizamos a través de las computadoras o con ayuda de ellas, simplificándolas y por ende, nos preguntamos ¿cómo las computadoras procesan la información?, ¿cómo es que una máquina puede llegar a tener inteligencia asemejando el proceso del pensamiento humano?

Aún se está lejos de afirmar que las computadoras pueden substituir por completo el trabajo que realiza el sistema nervioso parte de la maravillosa maquinaria humana en el procesamiento de la información, sin embargo podemos decir que de manera parcial realizan algunos procesos que imitan a los del cerebro.

La inteligencia artificial “IA” es una rama de las ciencias computacionales que contesta estos cuestionamientos. Para comprender su aplicación, es necesario revisar los siguientes conceptos:

- Inteligencia.- Capacidad de entender o comprender y de resolver problemas (RAE, 2014).

- Algoritmo.- Está definido como un conjunto de pasos, procedimientos o acciones que nos permiten alcanzar un resultado o resolver un problema (Cairó, 1995). Los algoritmos se caracterizan por lo siguiente:
 - ✓ Los pasos a seguir deben ser precisos.
 - ✓ Dado un conjunto de datos de entrada idénticos, siempre arrojará los mismos resultados.
 - ✓ Tendrá un número de pasos finito.

El ejemplo típico por ser el que ilustra mejor a un algoritmo son las recetas de cocina, que son un conjunto contado de pasos (finitud) que indican a detalle (precisión) como preparar un guiso y que si se obedece al pie de la letra, siempre obtendremos como resultado un platillo sabroso y bien elaborado (determinismo).

- IA.- Es la simulación de la inteligencia humana en una máquina, para hacer que la máquina sea eficiente e identifique y use una parte del conocimiento en una serie de pasos para resolver un problema (Konar, 2000).

Al revisar estos conceptos, entendemos que todos los sistemas informáticos utilizan algoritmos, sin embargo los sistemas informáticos basados en inteligencia artificial se diferencian de los tradicionales por el hecho de simular la inteligencia humana, bajo este paradigma, los algoritmos utilizados en la inteligencia artificial se clasifican de acuerdo a la técnica que permita al sistema desarrollar dicha simulación.

La prueba de Turing, propuesta por Alan Turing (1950), se diseñó para probar una definición operacional y satisfactoria de inteligencia. Turing sugiere una prueba basada en la incapacidad de diferenciar entre identidades indiscutibles y seres humanos. Hoy en día, se puede decir que una computadora que supera la prueba es realmente inteligente. Según Turing la computadora debería de tener las siguientes capacidades:

- a) Procesamiento de lenguaje natural. Comunicación satisfactoria.
- b) Representación del conocimiento. Almacenamiento de lo que conoce y siente.

- c) Razonamiento automático. Utiliza razonamiento de la información almacenada para responder preguntas y extraer nuevas conclusiones.
- d) Aprendizaje automático. Adaptarse a nuevas circunstancias y para detectar y extrapolar patrones.

Por tanto, entre las principales áreas de aplicación de la IA se encuentran: el razonamiento, el aprendizaje, la planeación, la percepción, adquisición de conocimiento, y búsqueda inteligente. Pero, ¿cómo funcionan los sistemas basados en IA?

Los sistemas inteligentes o conocidos también como expertos, constan de una **base de conocimiento, una base de datos y un motor de inferencia.**

Así como las personas captan información a través de los sentidos, los sistemas de información deben ser alimentados con datos, entre los cuales, surgirán las variables para desarrollar los algoritmos. En sustitución de los sentidos, en las computadoras la información se suministra por medio de dispositivos periféricos (hardware) como el teclado, el micrófono y las cámaras por mencionar los más comunes. La información obtenida se almacena en **la base de datos.**

La base de conocimiento es un conjunto de datos que se obtienen de la base de datos, éstos constituyen las variables de los algoritmos y que de acuerdo al área de aplicación y técnica de IA, son utilizados para formar reglas, hacer inferencias, cálculos matemáticos o bien para ser utilizados como valores probabilísticos.

El motor de inferencia se refiere al funcionamiento de los algoritmos con la base de conocimiento para conducir a un resultado.

A la secuencia de expresiones (pasos) de los algoritmos se les conoce como sentencias y éstas, son escritas con pseudocódigo (código auxiliar en la descripción de un algoritmo para

que pueda leerse). Los lenguajes de programación son los que harán la traducción de las sentencias a sistema binario (ceros y unos) para que puedan ser leídas por la computadora.

Para aterrizar el concepto de IA y su funcionamiento, se presenta como desarrollo de este trabajo, un ejemplo de aplicación que consiste en un sistema de detección de riesgo escolar para el cual, se probaron tres algoritmos:

- 1) El KNN (algoritmo de agrupamiento y clasificación).
- 2) El Naive Bayes (algoritmo probabilístico).
- 3) El ID3 (algoritmo de inferencia por árbol de decisión).

Cabe mencionar que parte fundamental de este sistema fue el diseño y aplicación de un cuestionario como medio de captura de datos para alimentar al sistema (Gameplayart, 2008).

Detección de riesgo escolar de los alumnos de nuevo ingreso a la Escuela Superior de Cómputo (ESCOM) del Instituto Politécnico Nacional

Un proceso educativo se basa en la secuencia de pasos para la transmisión de saberes y valores. El modelo más simple del proceso educativo es el de una persona (la cual puede ser un docente, una autoridad, un padre de familia, entre otras) que se encarga de transmitir dichos conocimientos a otra u otras personas. Por lo tanto, existe un sujeto que enseña y otro(s) que aprende(n).

Sin embargo, la realidad suele ser mucho más compleja. El aprendizaje tiene matices que se ven afectados por diversos factores que pueden potenciar u obstaculizar el aprendizaje. Un aspecto de relevancia para fundar un adecuado proceso educativo, es adelantarse a los posibles problemas que pueda presentar la enseñanza y aprendizaje de un alumno en particular. Para ello, una detección temprana del posible riesgo académico de un alumno puede beneficiar en la prevención anticipada del potencial fracaso escolar de dicho alumno,

dando la oportunidad de planear acciones que permitan la eliminación, o en el peor de los casos la atenuación, del riesgo escolar del alumno.

Los procesos educativos pueden clasificarse en dos niveles: formal e informal. El nivel formal se desarrolla en instituciones educativas como escuelas o universidades, contando con docentes profesionales, programas de estudio aprobados y sistemas de evaluación que exigen al alumno el cumplimiento de ciertos objetivos.

El nivel informal puede desarrollarse en el seno del hogar, en la calle o incluso de manera autodidacta. Los conocimientos que asimilan quienes aprenden, en este caso, no están sistematizados. El proceso educativo incluso puede desarrollarse a distancia, sin que las personas involucradas en el mismo estén cara a cara o tengan un contacto personal.

Es justamente en el nivel formal, donde es posible sistematizar el proceso educativo, permitiendo lograr un aprendizaje dirigido a través de acciones concretas que apoyen en el mejor cumplimiento del proceso.

Como parte de estas acciones es posible implementar la detección temprana del riesgo escolar (O'Byrne, 2009; Kyoshaba, 2009), esta detección temprana debe realizarse una vez que el alumno ingresa a la institución educativa, de forma tal que pueda generarse un plan de intervención para la eliminación o atenuación del riesgo detectado (Mushtaq, 2012; Fraser, 2003). Realizar esta detección sería una labor titánica si se llevase a cabo de forma manual, sin embargo en la actualidad los avances tecnológicos permiten la automatización de procesos mediante el uso de la computación. Incluso aquellos procesos de índole cognitivo, propios del ser humano, pueden ser simulados a través de la computación.

La inteligencia artificial, como parte de las ciencias computacionales, puede servir como herramienta para aproximarse al razonamiento humano, y de esta forma permitir la automatización masiva de análisis complejos, como lo es la detección de riesgo escolar. En los siguientes apartados se describe como se ha aplicado la inteligencia artificial para la

detección de riesgo escolar de los alumnos de nuevo ingreso a la Escuela Superior de Computo (ESCOM) del Instituto Politécnico Nacional, con el afán de sistematizar y apoyar en el mejoramiento del proceso educativo realizado en la ESCOM.

Estrategia de detección de riesgo escolar

La estrategia que se llevó a cabo para la detección de riesgo escolar de los alumnos de nuevo ingreso a la ESCOM, consiste de cuatro acciones:

1. Elaboración de un instrumento para recabar información socio-económica y académica de cada uno de los alumnos.
2. Diseño e implementación de un recurso de Internet para la implantación del instrumento elaborado.
3. Diseño e implementación de tres algoritmos de inteligencia artificial para la detección de riesgo escolar de cada uno de los alumnos.
4. Generación de reportes individuales así como de un reporte global acerca de la detección de riesgo escolar obtenida de los alumnos.

Elaboración de un instrumento para recabar información socio-económica y académica de cada uno de los alumnos

El instrumento propuesto para recabar información socio-económica y académica de cada uno de los alumnos de nuevo ingreso es un cuestionario que se divide en siete secciones, cada una de las cuales atiende aspectos personales, socio-económicos y académicos del alumno. Las preguntas incluidas en el cuestionario son el resultado de un estudio previo, el cual fue realizado con la finalidad de detectar aquellas preguntas que potencialmente podrían apoyar en la mejor detección del riesgo escolar de un alumno. En este estudio se consultaron diversos instrumentos empleados a nivel institucional (PRONABES, 2012) y en otras universidades (UAEM, 2011; UNAM, 2011) para la recolección de datos socio-económicos de los alumnos.

Las secciones en que se divide el cuestionario utilizado en el presente trabajo son siete, siendo éstas las siguientes:

- a) Registro del alumno. En esta parte del cuestionario se pide al alumno que ingrese los datos generales relevantes para su registro. En la figura 1 se muestra la estructura de esta parte.

Alumno

148.204.57.111/cuestionario/Alumno.html

Registro del alumno

Proporcione la información solicitada en los siguientes campos.

1.- Boleta:

2.- Apellido paterno:

3.- Apellido materno:

4.- Nombre:

5.- Grupo:

6.- Edad:

7.- Género:

8.- Correo electrónico:

9.- Estado civil:

10.- Número de hijos:

Aceptar

b) Antecedentes académicos. En esta parte del cuestionario se pide al alumno que ingrese información académica relevante para el objeto de estudio. En la figura 2 se muestra la estructura de esta parte.

Antecedentes académicos

Proporciona la información solicitada en los siguientes campos.

1.- Institución de procedencia:

2.- Promedio de tu bachillerato:
 Solo con números y dos dígitos decimales, ejemplo: 8.20

3.- ¿Reprobaste materias?

4.- ¿Tuviste beca?

5.- ¿Recibiste orientación vocacional?

6.- ¿ESCOM fue tu primera opción en el examen de admisión?

7.- ¿Consideras que tienes los conocimientos necesarios en física y matemáticas para cursar esta carrera?

Figura 2. Antecedentes académicos del alumno.

Datos personales

Proporciona la información solicitada en los siguientes campos.

- 1.- ¿Quién sostiene tus estudios?
- 2.- ¿Cuántos hermanos tienes?
- 3.- ¿Cuántas personas dependen de tus padres o tutor?
- 4.- ¿Cuál es el ingreso mensual aproximado de tu familia?
- 5.- ¿Trabajas?
- 6.- ¿Dependen personas de ti económicamente?
- 7.- En promedio, ¿cuánto gastas mensualmente al asistir a la escuela?
- 8.- ¿Con quién vives?
- 9.- ¿Eres derechohabiente de alguna institución de salud?

Aceptar

Figura 3. Datos personales del alumno.

- c) Datos personales. En esta parte del cuestionario se pide al alumno que ingrese los datos personales relevantes para el objeto de estudio. En la figura 3 se muestra la estructura de esta parte.
- d) domicilio. En esta parte del cuestionario se pide al alumno que ingrese la información relacionada con su domicilio. En la figura 4 se muestra la estructura de esta parte.

Figura 4. Domicilio del alumno.

Figura 5. Hábitos y estudio.

e) Hábitos y estudio. En esta parte del cuestionario se pide al alumno que ingrese la información relacionada con sus hábitos cotidianos y de estudio. En la figura 5 se muestra la estructura de esta parte.

- f) Vivienda. En esta parte del cuestionario se pide al alumno que ingrese la información relacionada con la vivienda que habita el alumno. En la figura 6 se muestra la estructura de esta parte.
- g) Datos del padre o tutor y vocación. En esta parte del cuestionario se pide al alumno que ingrese la información relacionada con su padre o tutor, así como aspectos de su vocación. En la figura 7 se muestra la estructura de esta parte.

The screenshot shows a web browser window with the address bar displaying '148.204.57.111/cuestionario/Vivienda.php'. Below the browser, there is a red header with the text 'Vivienda'. The main content area is a yellow box containing the following text and form elements:

Proporciona la información solicitada en los siguientes campos.

- 1.- Tipo de vivienda: [dropdown menu]
- 2.- ¿Paga renta? [dropdown menu]
- 3.- ¿Actualmente vives con tu familia? [dropdown menu]
- 4.- ¿Con cuántas personas vives? [dropdown menu]
- 5.- ¿En el lugar que habitas hay un espacio específico para estudiar? [dropdown menu]

At the bottom of the yellow box is a button labeled 'Aceptar'.

Figura 6. Vivienda.

The screenshot shows a web browser window with the address bar displaying '148.204.57.111/cuestionario/Servicios.php'. Below the browser, there is a red header with the text 'Datos del padre o tutor y tu vocación'. The main content area is a yellow box containing the following text and form elements:

Proporciona la información solicitada en los siguientes campos.

- 1.- Indica el nivel máximo de estudios de tu padre o de la persona que funja como Jefe de familia. [dropdown menu]
- 2.- Indica la ocupación que es o fue del Jefe de familia. [dropdown menu]
- 3.- ¿Tienes hermanos cursando alguna carrera de nivel superior? [dropdown menu]
- 4.- ¿Cuánto tiempo haces diariamente para trasladarte del lugar donde habitas a la escuela? [dropdown menu]
- 5.- ¿Cuál es la principal razón por la que elegiste tu carrera? [dropdown menu]

At the bottom of the yellow box is a button labeled 'Aceptar'.

Figura 7. Datos del padre o tutor y vocación.

Diseño e implementación de un recurso de Internet para la implantación del instrumento elaborado

La arquitectura general del sistema propuesto para la detección de riesgo escolar se muestra en la figura 8. Esta arquitectura consta de dos elementos principales dentro de la capa de presentación: un recurso de Internet (interfaz en guinda) y una aplicación de escritorio (interfaz en azul). El recurso de Internet mantiene el cuestionario que se aplica a los alumnos de nuevo ingreso de la ESCOM, este recurso utiliza una base de datos dentro de la capa de datos, donde se almacena, de forma persistente, la información capturada a través de las siete secciones que conforma el cuestionario descrito en el apartado anterior.

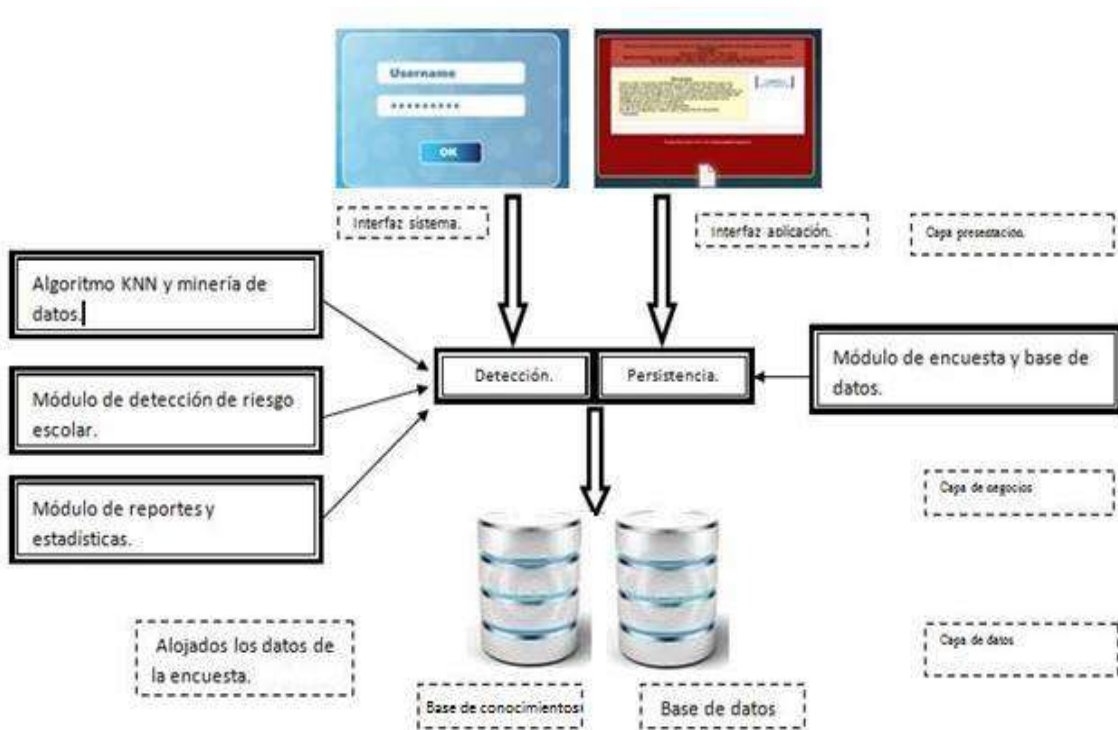


Figura 8. Arquitectura del sistema para la detección de riesgo escolar de los alumnos de nuevo ingreso a la ESCOM.

El recurso de Internet se ha diseñado e implementado utilizando herramientas computacionales tales como los lenguajes HTML, PHP (PHP, 2013) y JavaScript para su programación, así como el servidor Web Apache y el gestor de base de datos MySQL (MySQL, 2003), usando como sistema operativo a Linux. Todas estas herramientas son de uso libre, lo que significa que no existe ningún tipo de costo económico por las licencias para su uso. De la figura 1 a la 7 puede observarse la implementación del recurso de Internet para la implantación del cuestionario.

Diseño e implementación de tres algoritmos de inteligencia artificial para la detección de riesgo escolar de cada uno de los alumnos

La inteligencia artificial, como se mencionó en la introducción, es un área de las ciencias computacionales que pretende asemejar el razonamiento humano a través de algoritmos. Dentro de la inteligencia artificial se encuentra la minería de datos, cuyo objetivo fundamental es aprovechar el valor de la información focalizada y usar los patrones establecidos para obtener un mejor conocimiento de los datos, y de esta forma tomar decisiones más concretas y confiables.

Observando este objetivo, es como se decide utilizar la minería de datos para tratar la información recabada de los alumnos de nuevo ingreso para que a partir de ésta, se establezcan patrones que nos permitan detectar el riesgo escolar de un alumno mediante la aplicación de algoritmos de inteligencia artificial, los cuales permitan identificar y clasificar a cada alumno en relación a su potencial riesgo escolar.

Los algoritmos utilizados para la detección de riesgo escolar (capa de negocios de la figura 8) son tres: el algoritmo KNN, el algoritmo Naive-Bayes y el algoritmo ID3. KNN es un algoritmo de clasificación que almacena todos los casos disponibles y clasifica los nuevos casos en base a una medida de similitud, es decir, un nuevo caso que tan similar es a alguno que ya se tiene clasificado. Naive-Bayes es un algoritmo probabilístico, el cual indica la probabilidad de que un evento suceda, y esta probabilidad está definida por función. ID3 es un algoritmo basado en un árbol de decisión, un árbol de decisión es el resultado de un

proceso de clasificación cuya fuente de datos es llevada dentro de un árbol de predicción que representará un conjunto de reglas de decisión.

Cada uno de los anteriores algoritmos es implementado en una aplicación de escritorio programada en el lenguaje de programación Python (Python, 2003). Mediante esta aplicación se captura el número de boleta del alumno del cual se desea detectar su riesgo escolar, generándose como resultado que tipo de riesgo presenta de acuerdo al razonamiento de clasificación que realiza el algoritmo de inteligencia artificial. Adicionalmente, la propia aplicación genera recomendaciones para disminuir el riesgo detectado, haciendo uso del conocimiento que se encuentra en una base de conocimientos, la cual se construye a partir de los datos que se tienen almacenados de cada alumno. La figura 9 muestra parte de la interfaz de la aplicación.

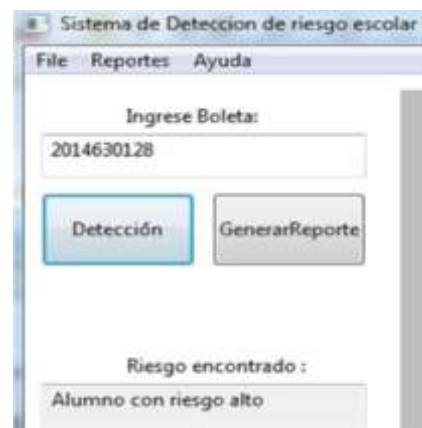


Figura 9. Interfaz de la aplicación del sistema de detección de riesgo escolar.

Generación de reportes individuales así como de un reporte global acerca de la detección de riesgo escolar obtenida de los alumnos

El sistema de detección de riesgo escolar permite como parte de su funcionalidad generar reportes individuales de cada alumno que fue analizado su riesgo, en este reporte se muestran los valores de las variables que fueron analizadas por cada uno de los algoritmos de inteligencia artificial descritos en el apartado anterior. Mediante este reporte personalizado es posible observar las características del alumno que lo clasifican bajo

alguno de los riesgos posibles (bajo, medio o alto). Con ello, es posible apoyar a la toma de decisiones por parte de las autoridades académicas del plantel para eliminar o atenuar el potencial riesgo que un alumno en particular pueda tener. Dentro del mismo reporte, el sistema muestra las recomendaciones que automáticamente genera como parte de ejecución del algoritmo. La figura 10 muestra un ejemplo de un reporte individual generado. También el sistema genera un reporte global de todos los alumnos detectados con su riesgo correspondiente.

Resultados obtenidos

El sistema de detección de riesgo escolar de los alumnos de nuevo ingreso a la ESCOM ha sido utilizado a partir del semestre escolar agosto-diciembre de 2013, la población total de alumnos de nuevo ingreso en ese semestre fue de 220, aplicándose el cuestionario a 104 de ellos. Estos 104 alumnos se usaron para generar la base de conocimientos que los algoritmos de inteligencia artificial utilizan. Durante el semestre actual, enero-junio, se aplicó el cuestionario a 60 alumnos de nuevo ingreso que se consideran como el grupo de prueba para la detección de riesgo escolar por parte del sistema. Después de la detección de riesgo por parte de los algoritmos sobre este grupo de prueba, se han obtenido los resultados mostrados en la tabla 1.

DREENIE

Sistema para la Detección de Riesgo Escolar de los Estudiantes de Nuevo Ingreso de la ESCOM
(DREENIE)

Boleta: 14038418
Nombre del alumno: yonatan
Correo del alumno:
Tipo de Riesgo: Riesgo ALTO
Recomendación: Tutoría, beca.

Bachillerato: CECYT
Promedio : 7
Materias Reprobadas: 4
Beca Anterior: SÍ
Orientación vocacional: SÍ
Quien sostiene tus estudios: AMBOS
Personas que dependen del padre : 3
Ingreso mensual: 4999
Tiempo de Estudio: 5
Espacio de Estudio: NO
Trabaja: NO
ESCOM primera opción: SÍ
Conocimiento FisMat: SÍ
Seguridad Social: IMSS

Figura 10. Reporte individual de la detección de riesgo escolar de un alumno.

Algoritmo empleado	Porcentaje de detección de riesgo escolar acertado
KNN	80.0
Naive-Bayes	80.0
ID3	70.7

Tabla 1. Resultados de la detección de riesgo escolar por parte del sistema.

Como puede observarse en la tabla 1 tanto el algoritmo KNN como Naive-Bayes detectan en un alto porcentaje de certidumbre, mientras que el algoritmo ID3 detecta con un menor porcentaje, sin embargo sigue siendo un porcentaje aceptable. Todo lo anterior indica que el sistema está muy próximo a un funcionamiento óptimo y con alentadores resultados para la detección de riesgo escolar de los alumnos. Esto permitirá anticiparse a potenciales problemas de fracaso escolar, permitiendo la sistematización y posible mejora del proceso educativo en la ESCOM.

Conclusiones

De los tres algoritmos se eligió el algoritmo KNN para implementar el sistema de detección de riesgo escolar de los alumnos de nuevo ingreso a la ESCOM por tener una eficiencia cercana al 80%, teniendo una incidencia del 20% entre falsos positivos. Además, este algoritmo reporta tres niveles de clasificación de riesgo (alto, medio y bajo), a diferencia de los algoritmos Naive Bayes e ID3 que reportan solo dos niveles (alto y bajo).

En los resultados de la aplicación del algoritmo KNN, es importante mencionar que la generación encuestada, en su mayoría obtuvo una clasificación en riesgo “alto”, y que variables como “Quién sostiene tus estudios” se encuentra con una relevancia importante, debido a que los alumnos etiquetados como exitosos reportan en la mayoría de los casos como valor de la variable mencionada, que es el padre quien sostiene los estudios; por lo

que se puede comprobar que el factor económico influye de manera positiva o negativa en el desempeño de un estudiante, infiriendo que quienes reciben el apoyo económico de los padres, principalmente del padre, dedican más tiempo a estudiar porque no tienen la necesidad de trabajar.

Es importante aclarar que la base de conocimiento es una base inicial, y que se pueden ir integrando nuevos elementos para tener más ejemplos de datos patrones, en nuestro caso, la base se construyó con datos del primer cuestionario siendo en su mayoría alumnos de riesgo escolar alto, así pues se tenía una alta incidencia de alumnos en riesgo (muchos patrones ejemplo en riesgo de este tipo), y pocos ejemplos de patrones de alumnos exitosos, lo que causó el 20% de falsos positivos.

También es necesario enfatizar que una parte medular en este tipo de sistemas, lo constituye la base de conocimiento que se construye a partir de la información (datos) que alimentan al sistema. En el caso del sistema presentado, el cuestionario fue el instrumento para recabar la información, siendo depurado después de la primera aplicación y de probarse con los algoritmos, dejando únicamente las preguntas que generaron respuestas (datos) útiles para la construcción de la base de conocimiento. Lo anterior fue con la finalidad de que en la segunda aplicación del cuestionario, éste fuera más ágil al responder, y se enfocará a obtener datos concretos y fiables para depurar a su vez la base de conocimiento.

Además de la depuración del cuestionario, se solicitó al Departamento de Gestión Escolar de la ESCOM, el promedio global que obtuvieron los alumnos al finalizar el semestre para comparar los resultados obtenidos en relación a los promedios y verificar si en verdad el algoritmo estaba arrojando datos satisfactorios.

Asimismo, cabe mencionar que debido al tiempo que se tenía para desarrollar el sistema, el cuestionario sólo se diseñó para obtener información sobre antecedentes académicos y factores socioeconómicos, no obstante se pueden diseñar nuevos cuestionarios que

contemplan otro tipo de factores como son los de tipo emocional, que influyen directamente en el desempeño académico, esto con el fin de crear nuevas bases de conocimiento y obtener reportes de detección de riesgo escolar en base a estos nuevos factores.

Por último, no podemos dejar a un lado que la veracidad con la que los alumnos responden el cuestionario podría afectar, aunque en un porcentaje mínimo, el resultado de la detección. Este aspecto no depende del funcionamiento del sistema sino de la forma en que se diseñe el cuestionario y los factores contextuales existentes en el momento de la aplicación, por lo que es recomendable que para la elaboración de este tipo de instrumentos que tienen el propósito específico de recabar información, se dedique tiempo, se busque asesoría de los expertos así como el apoyo de referencias confiables.

Bibliografía

Cairó, O. (1995). *Metodología de la programación, Algoritmos, Diagramas de Flujo y Programas, tomo1*. México: Alfaomega.

Fraser, W. J. y Killen, R. (2003). Factors influencing academic success or failure of first-year and senior university students: do education students and lecturers perceive things differently? *South African Journal of Education*, 23(4), 254 – 260.

Gameplayart. (2008). *Inteligencia Artificial: Descripción y Clasificación*. Disponible en: <http://www.gameplayart.com/learning/inteligencia-artificial-2n-descripcion-y-clasificacion/>

1.1

1.2 Globalemotion. (2011). *How the brain processes information*. Disponible en: <http://www.globalemotion.de/ted-wie-verarbeitet-das-gehirn-informationen-von-tom-wujec.html>

2.

3. Konar, A. (2000). *Artificial Intelligence and Soft Computing: Behavioral and Cognitive Modeling of the Human Brain*. EUA: CRC Press.

4.

5. Kyoshaba, M. (2009). *Factors affecting academic performance of undergraduate students at Uganda Christian University*. Uganda: Makerere University.

Medphys. (2010). *Basic Anatomy and Physiology of the Human Brain*. Disponible en: http://www.medphys.ucl.ac.uk/research/borg/homepages/florian/thesis/pdf_files/p25_34.pdf

Mushtaq, I. y Shabana, N. K. (2012, June). Factors affecting students' academic performance. *Global Journal of Management and Business Research*, 12(9), 16-22.

Neurocritic. (2012). [Does the Human Dorsal Stream Really Process Elongated Vegetables?](http://neurocritic.blogspot.mx/2012/03/does-human-dorsal-stream-really-process.html) Disponible en: <http://neurocritic.blogspot.mx/2012/03/does-human-dorsal-stream-really-process.html>

MySQL. (2003). *MySQL 5.0 Reference Manual*. Disponible en: <http://dev.mysql.com/doc/refman/5.0/es/index.html>

O'Byrne, J.W. y Britton, S. (2009, November). Using academic predictors to identify first year science students at risk of failing. *CAL-laborate International*, 17(1), 15-25.

PHP. (2013). *PHP Manual*. Disponible en: <http://us3.php.net/tut.php>

PRONABES. (2012). *Estudio socioeconómico*. Beca PRONABES. Disponible en: <http://uaca.reduaz.mx/descarga/pronabes/estudio.pdf>

Python. (2003). *Python Documentation*. Disponible en: <http://www.python.org/doc/>

RAE. (2014). *Diccionario Real Academia Española*. Disponible en: <http://www.rae.es/>

UAEM. (2011). *Estudio socioeconómico*. Universidad Autónoma del Estado de México. Disponible en: http://desarrollo.uaemex.mx/becas/Inc/Estudio_Socioeconomico1.pdf

UNAM. (2011). *Perfil de aspirantes y asignados a bachillerato y licenciatura de la UNAM*. UNAM. Disponible en: <http://www.planeacion.unam.mx/Publicaciones/pdf/perfiles/aspirantes/asp2010-2011.pdf>