

BOTBLOQ: Ecosistema integral para el diseño, fabricación y programación de robots DIY

Proyecto Financiado por el Centro de Desarrollo Tecnológico
Industrial (CDTI)

EXPEDIENTE: IDI-20150289

Cofinanciado por el Fondo Europeo de Desarrollo Regional
(FEDER) a través
del Programa Operativo Plurirregional de Crecimiento
Inteligente 2014-2020

ACRÓNIMO DEL PROYECTO: BOTBLOQ



Centro para el
Desarrollo
Tecnológico
Industrial



UNIÓN EUROPEA

Fondo Europeo de
Desarrollo Regional (FEDER)

Una manera de hacer Europa

E.6.1.4: Propuesta de Sistema Tutor Inteligente Adaptativo basado en Técnicas de Lógica Borrosa

RESUMEN DEL DOCUMENTO

BOTBLOQ: Ecosistema integral para el diseño, fabricación y programación de robots DIY

En este documento se presenta una propuesta para la utilización de técnicas sofisticadas de inteligencia artificial dentro de los módulos básicos que componen un Sistema Tutor Inteligente (STI). De esta manera se incorporan la utilización de herramientas como los Prototipos Deformables borrosos, las reglas borrosas y diversas técnicas de Minería de Datos con el fin de mejorar el funcionamiento de los módulos del tutor y del estudiante y ofrecer así rutas de aprendizaje personalizadas y adaptativas.

Índice

1. INTRODUCCIÓN	3
2. ANTECEDENTES	4
2.1. PROTOTOTIPOS DEFORMABLES BORROSOS	4
2.2. DESCUBRIMIENTO DE CONOCIMIENTO PROTOTÍPICO BORROSO	5
3. PROPUESTA	7
3.1. DISEÑO DE LOS CURSOS	7
3.2. EXTRACCIÓN DE CONOCIMIENTO DE UN LOG DE UN SISTEMA TUTOR	9
3.3. RUTA DE APRENDIZAJE SEGÚN LA TIPOLOGIA DEL ESTUDIANTE	11
3.4. RUTA DE APRENDIZAJE ADAPTATIVA	13
4. CONCLUSIONES	14
REFERENCIAS	15

1. INTRODUCCIÓN

Se plantea la creación de un Sistema de Tutorización Inteligente (en adelante ITS) que actúe como un tutor particular del estudiante y que, como un entrenador humano, posee libertad para actuar de acuerdo con las necesidades más complejas del estudiante. El ITS propuesto debe proveer de un modo de aprendizaje completamente adaptable conforme a los conocimientos previos y a la capacidad de evolución de cada estudiante.

La capacidad de adaptación del ITS al estudiante debe permitir analizar la evolución del mismo manejando un abanico de prototipos determinados por una serie de factores, es decir, que puede decidir qué tipo de evolución va a tener el estudiante adaptando a éste las diferentes formas de evolución básica que se definan. De esta manera se contará con una base de conocimiento con las estructuras básicas de evolución, se dispondrá de un mecanismo de equiparación entre la situación de cada estudiante y estas estructuras básicas y finalmente se generará una nueva definición de la evolución a partir de esta equiparación. Con este fin y como base científica del ITS se representará el conocimiento mediante prototipos y se utilizará la Lógica Borrosa para manejar la incertidumbre inherente al proceso.

Desde el punto de vista práctico se han utilizado técnicas de clustering basados en lógica borrosa que permiten el agrupamiento de datos de diferente naturaleza (cualitativos y cuantitativos) correspondientes a los atributos que definen a los tipos de alumnos (estilo de aprendizaje, notas, tiempos de estudio, rendimiento, etc.). Con la idea de diseñar un sistema tutor adaptativo rápido, con baja complejidad computacional, se ha optado por agrupar a los alumnos en función de sus características o atributos y así no tener que proponer tantas rutas de aprendizaje distintas como alumnos se tenga en el curso. Se han evaluado las técnicas realizadas no únicamente con las pruebas del sistema propuesto sino con datos públicos procedentes de Sistemas Tutores de prestigio como los ofrecidos en la KDD Cup 2010¹.

¹ <https://pslcdatashop.web.cmu.edu/KDDCup/>

2. ANTECEDENTES

En este apartado se presentarán los dos pilares científicos en la que se basa la propuesta del sistema tutor inteligente como son por un lado los prototipos deformables borrosos como herramienta de modelo de conocimiento y por otro las técnicas minería de datos utilizadas siguiendo el proceso de Descubrimiento de Conocimiento Prototípico Borroso.

2.1. PROTOTIPOS DEFORMABLES BORROSOS

La representación del conocimiento mediante Categorías Prototípicas Deformables Borrosas [1] también denominadas Prototipos Deformables Borrosos tiene sus orígenes en la teoría de Prototipos. La teoría de prototipos se desarrolló en el área de la Psicología Cognitiva con el fin de ofrecer un modelo acerca de la manera en que los humanos conceptualizamos la realidad. Investigaciones concluyeron que esta conceptualización la realizamos agrupando objetos en cuanto a sus propiedades o características. De esta manera, la teoría de prototipos propone una concepción de las categorías como clases heterogéneas, en las que habría algunos miembros más representativos de la categoría que otros, llamados prototipos. Rosch [2] y sus colaboradores conciben al prototipo como un ejemplar ideal, es decir, el elemento más representativo de la categoría, el que más rasgos comparte con los elementos de su propia categoría y menos con los de otras categorías. De manera que, este ejemplar ideal es posible compararlo con nuevos casos. De esta manera el prototipo tiene una doble caracterización: 1) Aquel elemento que es máximamente representativo para el concepto; 2) el elemento con máxima pertenencia al conjunto que representa al concepto.

Por otro lado, el profesor Zadeh [3] rechaza el concepto de prototipo procedente de la teoría clásica de prototipos argumentando que un solo objeto puede estar lejos de representar bien a un concepto, pero pertenecer completamente a él. Además, un prototipo casi nunca es un objeto sencillo sino un esquema borroso, que muestra en que medida distintos conceptos representan a un concepto. La prototypicalidad es entonces un problema de grado y no un solo elemento. Zadeh [3] propone una teoría de prototipos que se fundamenta en la teoría de conjuntos borrosos. En esta teoría, un concepto comprende un conjunto de prototipos que representan los diferentes grados de compatibilidad de las instancias con el concepto. De esta manera, las categorías no se conciben como clases discretas, es decir, con límites definidos, sino con límites borrosos, en los que se encontrarían los miembros periféricos de las categorías vecinas cuya transición sería gradual. Olivas [1] propuso una extensión de la teoría de prototipos de Zadeh a la que denominó Categorías Prototípicas Deformables Borrosas.

BOTBLOQ: Ecosistema integral para el diseño, fabricación y programación de robots DIY

Esta extensión consiste en:

- El número de categorías no está limitado a 3 como se especiada en la definición realizada por Zadeh.
- Una categoría no tiene porqué definir completamente a la nueva situación por lo que se definen mecanismos de adaptación a partir de la deformación de los prototipos definidos.
- Se hace uso del conocimiento experto, que proporciona una serie de ejemplos cuyas representaciones borrosas se realizan mediante números borrosos.

De esta manera, las Categorías Prototípicas Deformables Borrosas se basan en el concepto de prototipo para proporcionar información sintetizada de una colección de datos, con el fin de extraer conocimiento y patrones. Estos prototipos se obtienen a través del proceso de Descubrimiento de Conocimiento Prototípico Borroso (DCPB) [1], que es una extensión del KDD (Knowledge Discovery in Databases) [4] tal y como se verá en el siguiente apartado.

2.2. DESCUBRIMIENTO DE CONOCIMIENTO PROTOTÍPICO BORROSO

Olivas [1] con el fin de obtener comportamientos prototípicos a partir de grandes colecciones de datos propuso un método de extracción de conocimiento al que denominó Descubrimiento de Conocimiento Prototípico Borroso (DCPB). En el DCPB los prototipos se representan a través de marcos conceptuales definidos por sus correspondientes características. El DCPB surge como una extensión del KDD y hereda sus mismas etapas, excepto que en cada una de éstas interviene el conocimiento experto. La importancia de la intervención del conocimiento experto recae en que la extracción de conocimiento no es conducida por los atributos como en el KDD, sino por heurísticas que se obtienen en el preprocesado de los datos. En la Figura se observa el proceso del DCPB.



BOTBLOQ: Ecosistema integral para el diseño, fabricación y programación de robots DIY

A continuación, se describen brevemente cada una de las fases del DCPB:

- **Selección:** Con el fin de facilitar la manipulación de los datos con alta dimensionalidad, en esta etapa se obtiene un subconjunto reducido de atributos que contribuyen de manera significativa al proceso de extracción de conocimiento. Este subconjunto es posible obtenerlo a través de: 1) la experiencia del experto, 2) métodos de reducción de características existentes en la literatura [5, 6, 7]. En este documento se utiliza el término *tarjeta de datos* para dar nombre al subconjunto reducido tanto de atributos como de datos, sobre el cual se extraerá conocimiento.
- **Preproceso:** En esta fase se tratan los problemas que presentan los datos tales como: ausencia, incompletitud e inconsistencia. Otro problema se presenta cuando algunos datos no se ajustan al comportamiento general de los datos restantes, por lo que son llamados valores extremos (outliers) [8]. Los outliers puede ser datos correctos que son distintos a los demás.
- **Transformación:** En esta fase intervienen distintas técnicas que ajustan los datos al problema que se intenta resolver. Las tareas que se llevan a cabo en esta etapa son: convertir un tipo de dato a otro. Una de estas tareas, la normalización consiste en escalar los valores de un atributo, de tal manera que se reducen a un rango específico. La normalización es muy importante en algoritmos que usan medidas de distancia debido a que evitan que atributos con rangos grandes prevalezcan sobre atributos con rangos pequeños.
- **Minería de datos:** Esta etapa es una de las más importantes del proceso DCPB, ya que a partir del conjunto de datos bajo estudio se obtienen definiciones paramétricas en forma de marcos (tablas) con sus respectivas características (atributos). Una definición paramétrica se refiere a los resúmenes que describen al conjunto de datos bajo estudio. Estas definiciones paramétricas se pueden obtener a través de los distintos métodos que intervienen en esta etapa tales como: técnicas de agrupamiento (clustering), funciones de resumen, clasificación, entre otros. El modelo de datos generado en esta etapa estaría representado por los marcos, los cuales son una descripción de los patrones o las relaciones entre los datos, a partir del cual es posible obtener predicciones, comprender los datos o explicar situaciones pasadas.
- **Prototipos:** Los prototipos que se obtienen del proceso de Descubrimiento de Conocimiento Prototípico Borroso se representan como una matriz. Las filas representan a las características; las columnas los valores (mínimo, promedio y máximo, etc.), que pueden tomar dichas características. El uso de matrices para representar a los prototipos se debe a que al almacenar los datos como intervalos se toma en cuenta la variabilidad e incertidumbre en los mismos [9].

3. PROPUESTA

El Sistema Tutor Inteligente propuesto debe integrar todos los parámetros que definen el proceso de enseñanza aprendizaje los cuales se pueden englobar en dos conceptos:

- Ruta de Aprendizaje o Learning Path: secuencia de actividades que permite al estudiante lograr un determinado objetivo.
- Estilo o curva de aprendizaje (Learning Curve): que indica la evolución en el éxito del aprendizaje del estudiante a lo largo del tiempo.

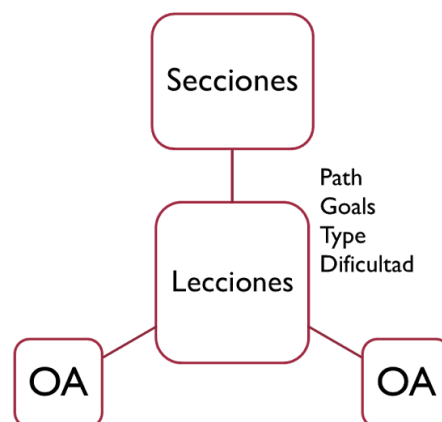
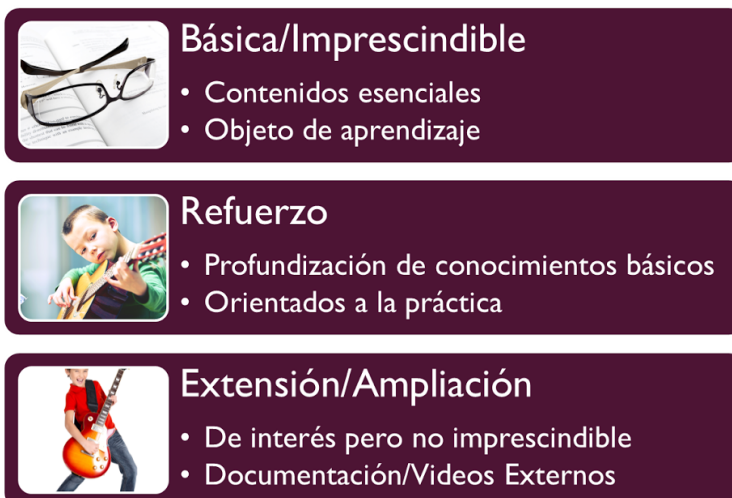
La base científica del sistema tutor residirá en disponer y manejar una base de conocimiento que contenga las diferentes curvas de aprendizaje que se pueden presentar dentro del sistema tutor y además pueda adaptar éstos a las características especiales de cada estudiante. Con este fin, se utilizará una representación de los diferentes estilos de aprendizaje en forma de prototipo complejo. Cada prototipo, corresponderá a un estilo de aprendizaje teórico y estará compuesto por los parámetros y actividades que definan su línea base y por una función de pertenencia que permita conocer la afinidad entre cada prototipo y la situación del estudiante para realizar posteriormente su adaptación de su ruta de aprendizaje.

3.1. DISEÑO DE LOS CURSOS

En primer lugar, se define de forma completa una estructuración de un curso de forma que dé lugar a las diferentes rutas de aprendizaje (secuencia de objetos de aprendizaje) personalizadas. De esta manera, los cursos van a estar estructurados en secciones y lecciones. Cada uno de estos elementos tendrá unos objetivos de aprendizaje especificados siguiendo la taxonomía de Bloom².

² <http://www.bloomstaxonomy.org/>

BOTBLOQ: Ecosistema integral para el diseño, fabricación y programación de robots DIY



En resumen, los cursos se van a organizar en una estructura jerárquica de 4 niveles:

- **Cursos:** El primer nivel de la jerarquía, posee la información general del curso, una lista con las secciones asignadas a él, una lista con sus objetivos principales y otros datos para el funcionamiento interno como el historial o las estadísticas.
- **Secciones:** Son el segundo nivel, puede haber más de una sección en un curso. También posee información general de la propia sección, otra lista de objetivos y una lista con las lecciones asignadas a ella.
- **Lecciones:** Son el tercer nivel de la jerarquía y el más importante de todos. Poseen información general sobre la lección, una lista de objetivos, la dificultad y el tipo de la actividad que representa y una lista con los Objetos de Aprendizaje. Con respecto al tipo hay que destacar que existen tres tipos de actividad: Básica, de Refuerzo y de Ampliación. Las actividades Básicas son las que todo alumno debe cursar, las que contienen el material esencial del curso y que se debe superar en todos los casos. Las actividades de Refuerzo consisten en lecciones (por lo general) más sencillas y que complementan los conocimientos de los estudiantes menos avanzados para una posterior mejora en sus resultados. Las de Ampliación están enfocadas a los estudiantes más cualificados y con mejores resultados, son lecciones con contenidos interesantes, pero no obligatorios ni esenciales.
- **Objetos de Aprendizaje:** para finalizar la jerarquía de los cursos está la lista con los Objetos de aprendizaje asignados a una lección. Por lo general todos los LOMs de una lección poseen la misma información, pero en distinto formato, de esta forma los estudiantes podrán cursar la actividad de la forma que mejor

BOTBLOQ: Ecosistema integral para el diseño, fabricación y programación de robots DIY

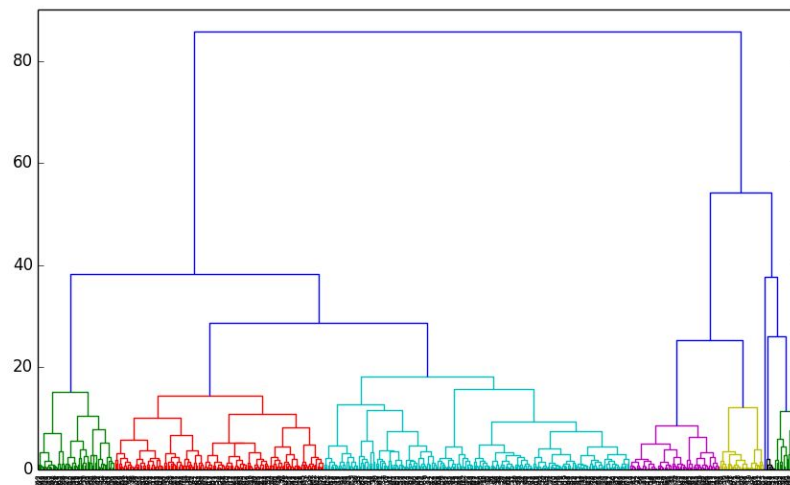
se adapte a su estilo de aprendizaje. Este formato puede ser video, audio, ejercicio o lección teórica.

El alumno no deberá de preocuparse por la estructura o por el orden en el que está realizado el curso, sino que el sistema se encargará de generar la ruta de aprendizaje personalizada conforme va avanzando el estudiante en dicho curso.

3.2. EXTRACCIÓN DE CONOCIMIENTO DE UN LOG DE UN SISTEMA TUTOR

A continuación, se detallará el proceso por el cual el sistema tutor incorpora funcionalidades basadas en conocimiento. Se trata de un proceso de minería de datos donde se han utilizado los datos facilitados por la KDD Cup 2010. Los pasos seguidos en este proceso son los siguientes

- **Selección:** El primer paso supone el estudio de la fuente de datos que consiste en el log de interacción de estudiantes con un sistema tutor inteligente. Cada interacción almacena información sobre el estudiante y sus resultados al resolver ejercicios matemáticos resueltos por el sistema. Se seleccionan las variables correspondientes y se preprocesan los valores erróneos y/o incompletos que existen en la base de datos. Posteriormente se construye la Tarjeta de Datos o conjunto de datos objetivo, que estará basado en los indicadores más relevantes para representar el rendimiento de un estudiante en el sistema tutor, por ejemplo, estos pueden ser: promedio de pistas solicitadas, duración en la realización de cada problema, porcentaje de problemas resueltos al primer intento, número de problemas cursados, etc.
- **Clustering:** A partir de estos datos se procederá a realizar un proceso de descubrimiento de conocimiento (data mining) que nos permitirá determinar los diferentes tipos de evolución que está presentando el estudiante en el sistema tutor. Se utilizarán algoritmos de agrupamiento (clustering) para en primer lugar disponer de una imagen lo más realista posible sobre los diferentes tipos de evolución que se están observando en los estudiantes. Estos resultados son utilizados para juzgar cuántos grupos de comportamientos distintos se pueden identificar entre los estudiantes. Cada uno de estos grupos representará a un prototipo provisional.



- **Caracterización de los prototipos:** Una vez obtenidos los grupos de alumnos se realizó un estudio para comprender que significaban cada uno de ellos y que características tenían en común los estudiantes de cada grupo. Con este fin se realizó una agrupación de los estudiantes según el resultado del clustering y se compararon sus datos. Además, se calcularon en cada grupo los valores máximos, mínimos y medios de cada uno de los indicadores.
 - **Grupos de Bajo Rendimiento (7,1,2):** Se consideran los grupos con peores resultados ya que con el mínimo de unidades, problemas y pasos obtiene también el mínimo en pasos correctos y el máximo en tiempo medio y número de pistas solicitadas.
 - **Grupos de Rendimiento Medio (3,4,6):** Entre los grupos 3, 4 y 6 arrojan unos resultados muy similares. Obtienen el mayor porcentaje de pasos resueltos correctamente y es el segundo con menor número de pistas solicitadas por problema. Destacan de forma menos significativa en promedio de problemas por unidad y tiempo medio por problema.
 - **Grupo de Rendimiento Alto:** El grupo 5 es considerado el mejor grupo. Es el segundo grupo con más unidades, problemas y pasos realizados, obteniendo además un alto porcentaje de pasos resueltos, el segundo mejor tiempo medio por problema y un alto número medio de habilidades necesarias por problema.
- **Extracción de reglas mediante árboles de decisión:** Mediante algoritmos de clasificación como los árboles de decisión se determinarán los factores que están siendo más relevantes para determinar el patrón de rendimiento de los estudiantes. Además, se extraerán las reglas necesarias con el fin de determinar que tipo de comportamiento ofrecerá cada estudiante en cada momento. Estas reglas son traducidas a un formato entendible por el sistema

BOTBLOQ: Ecosistema integral para el diseño, fabricación y programación de robots DIY

tutor y serán utilizadas por este de forma que el grupo que se le asigne al estudiante dependerá de su comportamiento previo en los cursos que haya realizado dentro del sistema tutor.

```
[{
  "condition": "function(R) { R.when((this.problems <= medium) &&
    (this.skills <=short) && (this.steps <=short));  ",
  "consequence": "function(R) {  this.group = 1;  R.stop();  "
}, {
```

3.3. RUTA DE APRENDIZAJE SEGÚN LA TIPOLOGIA DEL ESTUDIANTE

El proceso de enseñanza y aprendizaje del alumno estará completamente monitorizado mediante el análisis del log que generará el Sistema Tutor. El log del sistema tutor registrará información como: el identificador del estudiante, el nombre del problema resuelto, la unidad a la que pertenece el mismo, los nombres de los pasos en que se descompuso el problema, el tiempo empleado en resolver un paso, las ayudas adicionales (pistas) que el estudiante haya recibido del sistema, el número de veces que el estudiante regreso a resolver un mismo paso, las habilidades requeridas para resolver el paso, etc.

En el momento que un alumno se matricule en un curso se determina el grupo del estudiante a partir de la ejecución de las reglas de producción extraídas en el proceso de minería de datos y los valores de los indicadores que representan el comportamiento previo del estudiante como son los siguientes:

- Dificultad para resolver las actividades, viene dada por la capacidad que posee el estudiante para comprender/interpretar lo que se le pide
- Número de intentos realizados para resolver una actividad, se refiere al número de veces que el estudiante intentó resolver una actividad
- Tiempo empleado para resolver las actividades, se refiere al tiempo que le lleva a un estudiante para resolver una actividad
- Solicitud de ayuda adicional al sistema tutor, cuando no posee la habilidad para solucionar la actividad

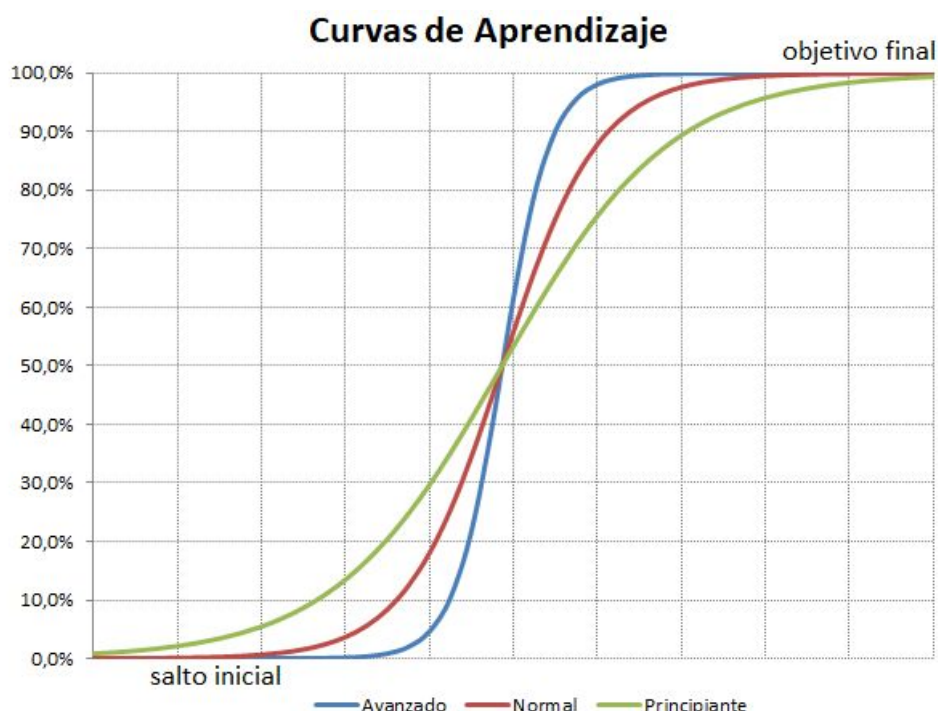
Además de obtener este dato, se obtiene el nivel de conocimiento que el estudiante tiene con respecto al curso en cuestión, para ello se comparan los objetivos logrados

BOTBLOQ: Ecosistema integral para el diseño, fabricación y programación de robots DIY

con los que pretende lograr y se establece el nivel (la lección) por la que debe comenzar, evitando así repeticiones innecesarias.

	Conocimientos Previos		
Capacidad	Altos - Largo	Básicos - Corto	No Existen - Sin salto
A	advanced	advanced	advanced
B	advanced	advanced	medium
C	advanced	medium	medium
D	medium	medium	medium
E	medium	medium	beginner
F	medium	beginner	beginner
G	medium	beginner	beginner

En base a estos dos valores se calcula el tipo de ruta de aprendizaje que va a seguir el estudiante durante el curso (Ver tabla anterior). Se determinan para ello tres tipos de estudiante que corresponde a tres rutas de aprendizaje distintas para realizar el curso, que garantizan llegar al mismo objetivo de aprendizaje final, pero con velocidades diferentes.



BOTBLOQ: Ecosistema integral para el diseño, fabricación y programación de robots DIY

De esta manera, al estudiante “avanzado” se le propondrá una ruta de aprendizaje corta de duración y de dificultad alta, mientras que al estudiante “principiante” se le recomendará una ruta de aprendizaje de larga duración y dificultad baja. El estudiante “medio” dispondrá de una ruta de aprendizaje promedio en duración y de dificultad intermedia.

En la siguiente tabla se puede observar las reglas que gobiernan la ruta de aprendizaje de un estudiante según la tipología que se calcula al matricularse en un curso.

Tipo de estudiante	Supera actividad básica	No supera actividad básica	Supera actividad refuerzo	No supera actividad refuerzo	Supera actividad ampliación	No supera actividad ampliación
Advanced	Refuerzo más difícil	Refuerzo más difícil	Básica siguiente	Otra por refuerzo	Básica siguiente	Repetir ampliación
Medium	Refuerzo aleatorio	Repetir la misma	Refuerzo siguiente (2 máx.)	Repetir la misma		
Beginner	Refuerzo más fácil	Repetir la misma	Refuerzo siguiente	Repetir básica anterior		

3.4. RUTA DE APRENDIZAJE ADAPTATIVA

Tras personalizar la ruta de aprendizaje a seguir en un curso se persigue conseguir un sistema tutor que permita adaptarse no solo al tipo de estudiante sino también a su situación dentro de cada curso. Esta capacidad de adaptación se logra con el análisis del comportamiento del estudiante durante el curso, análisis que ofrecerá como resultado una posible actualización de su ruta de aprendizaje según mejore o empeore sus resultados. En suma, el estudiante podrá actualizar su tipología (principiante, medio o avanzado) durante la realización del curso, de forma que pueda subir de nivel (de medio a avanzado o de principiante a medio) o bajar (de avanzado a normal o de normal a principiante).

Para realizar este análisis se tienen en cuenta dos indicadores principales:

- **Rendimiento del estudiante:** Se calcula a partir del porcentaje de actividades terminadas correctamente según su dificultad. De esta manera se considera un rendimiento alto cuando no se producen fallos en lecciones de dificultad normal o baja y ocasionalmente en las lecciones difíciles. Por el contrario, un rendimiento bajo supondrá que se producen fallos no ocasionales en lecciones de dificultad baja.
- **Tiempo de realización de las actividades:** Se calcula mediante la comparación con los tiempos medios de realización de actividades y determinando si el comportamiento del estudiante está dentro o fuera de control utilizando para ello la teoría de control básica³. De esta forma, si el estudiante en alguna actividad ha tardado un tiempo superior o inferior que el límite de control, establecido mediante la media y la desviación típica o bien se producen de forma consecutiva N (se fija como tamaño de la ventana) actividades con tiempo por encima de la media (tiempo alto) o por debajo de la media (tiempo bajo) se determina que el tiempo está fuera de control y que la ruta de aprendizaje pueda cambiar.

Una vez que se calculan estos indicadores y mediante la utilización de reglas basadas en etiquetas lingüísticas se establece cual debe ser la modificación, si procede, de la ruta de aprendizaje del estudiante.

³ <http://www.itl.nist.gov/div898/handbook/pmc/section3/pmc322.htm>

	Rendimiento		
Tiempo	Alto	Medio	Bajo
Bajo	Sube	Baja	Baja
Normal	Sube	Mantiene	Baja
Alto	Mantiene	Mantiene	Baja

Una vez puesto en marcha el sistema adaptativo hace alta una actualización de los tipos de estudiante ya que con solo tres no se reflejaban los casos extremos. Para solventar este problema se incluyeron dos tipos más, el **Principiante-** y el **Avanzado+**. El tipo **Principiante-** describirá a los estudiantes con un rendimiento muy por debajo de la media, es decir, con grandes dificultades para obtener conocimientos y que necesitan un apoyo mayor que el resto. En cambio, el tipo **Avanzado+** se le asignará a los estudiantes con unos resultados sobresalientes y con una gran facilidad para obtener conocimientos con respecto al resto de sus compañeros.

4. CONCLUSIONES

Con el fin de determinar el comportamiento de la propuesta se hizo uso de algunas aproximaciones cuantitativas, que comúnmente son utilizadas para evaluar los resultados arrojados por los métodos de predicción utilizando para ellos los datos proporcionados por la KDD Cup 2010.

Las aproximaciones que se utilizaron en este trabajo fueron: el RMSE (Root Mean Squared Error), el MMRE (Mean Magnitude of Relative Error) y PRED(25), que resultan adecuados para predecir las notas de los estudiantes o lo bien que realizan un ejercicio.

Se ha podido observar que con la representación borrosa de prototipos se obtiene un nivel de error bajo, comparado con las aproximaciones basadas en líneas base. Los resultados que se obtuvieron con MdmRE y MMRE caen dentro del 15% de error aceptable, mientras que con PRED(25) más del 75% de los casos tuvieron predicciones dentro del 25% de sus valores reales. A su vez se ha detectado que existe una diferencia notable en los resultados arrojados por las diferentes aproximaciones, específicamente con los estudiantes con rendimiento académico Normal y Principiante, por lo que se puede concluir que el uso de Prototipos Deformables Borrosos resulta más conveniente. Cabe señalar que en algunas ocasiones los resultados que arroja el MdmRE son mejores que los que se obtienen con el MMRE,

BOTBLOQ: Ecosistema integral para el diseño, fabricación y programación de robots DIY

esto se debe a que las predicciones realizadas pueden fallar en momentos concretos de la evolución del rendimiento académico de los estudiantes.

Además de esta evaluación se mantendrá un control constante del funcionamiento del sistema con el fin de permitir su evolución y adaptación a las necesidades cambiantes de los alumnos. De esta manera, se realizará un análisis más profundo de los logs y encuestas que se realizan a los alumnos con el fin de verificar y controlar el buen funcionamiento del sistema tutor tanto desde el punto de vista técnico (rendimiento, usabilidad), como desde el punto de vista funcional (utilidad, atractivo de los contenidos, novedad, etc.). También se realizará un análisis sobre los resultados obtenidos por los alumnos comparándolos con los que teóricamente deberían haber conseguido al final del curso.

Referencias

1. Jose A. Olivas. Contribución a un estudio experimental de predicción basada en Categorías Deformables Borrosas. Tesis de doctorado, Universidad de Castilla La Mancha, 2000.
2. Eleanor Rosch. Cognitive representations of semantic categories. *Experimental Psychology*, 104(3):192–233, 1975
3. Lotfi A. Zadeh. A note on prototype theory and fuzzy sets. *Cognition*, 12(3):291 – 297, 1982.
4. Usama Fayyad, Gregory Piatetsky-Shapiro, y Padhraic Smyth. The KDD Process for Extracting Useful Knowledge from Volumes of Data. *Communication of the ACM*, 39(11):27–34, 1996
5. Isabelle Guyon y Andre Elisseeff. An Introduction to Variable and Feature Selection. *Machine Learning Research*, 46:389–422, 2003.
6. Huan Liu y Hiroshi Motoda. *Feature Selection for Knowledge Discovery and Data Mining*. Kluwer Academic Publishers, 1998.
7. Huan Liu y Hiroshi Motoda. *Computational Methods of Feature Selection*. Chapman and Hall/CRC, 2008
8. Jacob Cohen, Patricia Cohen, Stephen G. West, y Leonas S. Aiken. *Applied multiple regression/correlation analysis for the behavioral sciences*, tomo 1. Lawrence Erlbaum Associates, 200
9. Chenyi Hu y R. Baker Kearfott. Interval matrices in knowledge discovery. En *Knowledge Processing with Interval and Soft Computing*, *Advanced Information and Knowledge Processing*, págs. 1–19. Springer London, 2008.