

INTRODUCCIÓN A LAS TÉCNICAS DE APRENDIZAJE AUTOMÁTICO (PARTE I): ESTRUCTURACIÓN DE UN CURSO DOCENTE

Elisabet Golobardes, Ester Bernadó, Xavier Llorà, Josep M. Garrell, Maria Salamó, Joan Camps y David Vernet

Grupo de Investigación en Sistemas Inteligentes, Departamento de Informática, Ingeniería i Arquitectura La Salle, Universitat Ramon Llull, Passeig Bonanova 8, 08022 - Barcelona

e-mail: {elisabet, esterb, xevil, josepmg, mariasal, joanc, dave}@salleURL.edu

RESUMEN: En este artículo se presentan los Seminarios del Grupo de Investigación en Sistemas Inteligentes de *Enginyeria i Arquitectura La Salle* de la *Universitat Ramon Llull*, estructurados como un curso de introducción a las técnicas de aprendizaje automático. Este curso ofrece un marco común a los alumnos que cursan asignaturas de libre elección, o bien realizan sus trabajos y/o proyectos de final de carrera, dentro de nuestro grupo de investigación.

1.- INTRODUCCIÓN

Los Seminarios del Grupo de Investigación en Sistemas Inteligentes de Ingeniería i Arquitectura La Salle (EALS), de la Universitat Ramon Llull (URL) pretenden ser un curso que aglutina diferentes asignaturas de libre elección de la Ingeniería Técnica de Informática de Sistemas y de la Ingeniería en Informática. Al mismo tiempo, pretende establecer un marco común tanto los Trabajos de Final de Carrera como los Proyectos Final de Carrera realizados dentro del grupo. Finalmente, también incluye a los alumnos "colaboradores", en general, con el Grupo de Investigación en Sistemas Inteligentes.

Como se puede observar, el perfil y los conocimientos de los distintos alumnos agrupados bajo el paraguas de los Seminarios del Grupo de Investigación en Sistemas Inteligentes es muy variado. Así pues, para entender como se adecua el temario a cada tipo de estudiante, previamente presentaremos el trabajo que realiza el Grupo de Investigación en Sistemas Inteligentes, y a quién van dirigidos los seminarios dentro del contexto que definen los planes de estudio de EALS.

Una vez descrito el contexto de los Seminarios del grupo, pasaremos a presentar sus objetivos, el planteamiento diseñado, la planificación global de un curso, los criterios de evaluación y algunas conclusiones.

2.- GRUPO DE INVESTIGACIÓN EN SISTEMAS INTELIGENTES

El grupo de investigación en Sistemas Inteligentes (GrupoSI) se enmarca dentro del Grupo de Investigación en Inteligencia Artificial de la Universidad Ramon Llull. Las principales actividades del GrupoSI se basan en las distintas técnicas del paradigma del aprendizaje automático (o aprendizaje artificial). Las líneas básicas de trabajo discurren dentro de las familias del *aprendizaje evolutivo* (algoritmos genéticos, sistemas de clasificación, programación genética y vida artificial), del *aprendizaje analógico* (razonamiento basado en casos), del *aprendizaje conexionista* (redes neuronales), del *aprendizaje inductivo* (principalmente árboles de decisión), y de distintos *híbridos* entre las diferentes técnicas. Nuestro trabajo se ha centrado en resolver problemas de clasificación, de optimización y/o de diagnóstico. Concretamente, el trabajo de investigación del grupo se ha desarrollado, mayoritariamente, dentro del marco que ofrecen distintos proyectos subvencionados en los cuales estamos involucrados:

- Aprendizaje artificial para la predicción y la diagnosis (Funitec)
- Diagnóstico automática de cáncer de mama con técnicas de IA (Funitec)
- Generación de modelos acústicos de calles de una ciudad con técnicas de IA (Funitec)
- Generación automática de estrategias de demostración para sistemas paralelos (Funitec)
- Participación en la *Xarxa Temàtica d'Intel·ligència Artificial* de la Generalitat de Catalunya
- Implementación y estudio de herramientas de inteligencia artificial aplicadas a plataformas de enseñanza abierta a distancia y a las redes ATM que las soportan (CICYT/TEL98-0408)
- *Automatic classification of long bone fractures from X-Ray images using the Müller's CCF system* (Maurice E. Müller Foundation)
- Sistema de predicción del riesgo de cáncer de mama basado en técnicas de visión por computador y de aprendizaje artificial automático (FIS 00/0033-02)

3.- ¿A QUIÉN VAN DIRIGIDOS LOS SEMINARIOS?

Ingeniería i Arquitectura La Salle ofrece 7 titulaciones técnicas (una de propia: graduado técnico en multimedia) y 5 de superiores (una de propia: graduado superior en multimedia). Concretamente, se pretende que las titulaciones de la Ingeniería (6+4) tengan un primer curso muy común y que, al mismo tiempo, a partir de asignaturas optativas y de libre elección un estudiante de una titulación pueda adquirir fácilmente un perfil de otra titulación. Una de las maneras de ofrecer esta interdisciplinariedad, es a partir de unas asignaturas de libre elección (denominadas intensificación y especialización de cada sección de la Escuela) que el estudiante puede cursar “colaborando” con los distintos grupos de trabajo de los diferentes departamentos de EALS. Aunque la participación puede ser de alumnos de carreras distintas, los estudiantes que colaboran con nuestro grupo tienen un perfil mayoritariamente informático. Así pues, desde esta perspectiva, en nuestro grupo pueden “colaborar” estudiantes de diferentes titulaciones, principalmente cursando las asignaturas de libre elección, realizando su trabajo o proyecto de final de carrera (ver tablas 1 y 2)⁴⁸ o, simplemente, colaborando con nosotros.

⁴⁸ El GrupoSI desarrolla su actividad dentro del Departamento de Informática y, más concretamente, sus áreas de interés son afines a las asignaturas de la sección de Tecnología del Software. Por este motivo, las intensificaciones y especializaciones son en Tecnología del Software.

Tabla 1. Asignaturas correspondientes al plan de estudios de Ingeniería Técnica en Informática de Sistemas, EALS, URL

Asignatura	Tipo	Créditos
Intensificación en Tecnología del Software	Libre elección	9
Especialización en Tecnología del Software	Libre elección	9
Trabajo Final de Carrera (TFC)	Obligatoria	8

Tabla 2. Asignaturas correspondientes al plan de estudios de Ingeniería en Informática, EALS, URL

Asignatura	Tipo	Créditos
Intensificación en Tecnología del Software	Libre elección	9
Especialización en Tecnología del Software	Libre elección	9
Proyecto Final de Carrera (PFC)	Obligatoria	8

Con el propósito de agrupar todos estos estudiantes, desde el GrupoSI, elaboramos los *Seminarios* a partir de los cuáles pueden formarse en los principales temas en los que trabajamos y, así, pueden desarrollar sus créditos de libre elección, o bien su TFC o PFC, o bien, su colaboración. Pero estos Seminarios se tendrán que adecuar al perfil de cada estudiante.

4.- CONTENIDO DE LOS SEMINARIOS

Los Seminarios del Grupo SI se centran en las áreas siguientes: algoritmos genéticos (GA), redes neuronales (NN), programación genética (GP), razonamiento basado en casos (CBR) y sistemas clasificadores (CS). Concretamente, les damos una introducción sobre:

- *Computación Evolutiva* introduciendo tres de sus posibles vertientes para la clasificación automática: GA, GP y CS. También se les muestra alguna herramienta propia desarrollada dentro del GrupoSI para trabajar con GA [1,2,3,4,9].
- *Redes Neuronales*. Se pretende realizar una introducción de éstas, y explicar, básicamente, las más clásicas para poder entender el problema de la no linealidad: el Perceptrón, la Adaline, la Madaline y la Backpropagation[5]. Se les muestra alguna herramienta propia del GrupoSI, o bien, alguna herramienta de libre distribución (e.g. TRAJAN 2.0).
- *Razonamiento Basado en Casos*. Ofrecemos una breve introducción a la filosofía del CBR, y detallamos mejor el CBR encarado a resolver problemas de clasificación. Concretamente, les explicamos algunas soluciones que hemos dado para resolver el problema del diagnóstico de tejidos de glándulas mamarias [3,4,7], o bien, para el diagnóstico de mamografías a partir de microcalcificaciones [10,11,12].

En las exposiciones más prácticas procuramos que participen alumnos de otros años "antiguos" siguen colaborando con nuestro grupo.

5.- PLANTEAMIENTO DE LOS SEMINARIOS

Los Seminarios se han diseñado distinguiendo dos tipos de alumnos: (1) los colaboradores (colaboradores o alumnos que cursan las correspondientes asignaturas de libre elección); y (2) los proyectistas (alumnos que realizan su TFC o PFC). Concretamente, presentamos la estructura de los Seminarios para un alumno que trabaja con nosotros por primera vez⁴⁹. Estos Seminarios se estructuran en tres partes que, más o menos, corresponden a los trimestres naturales.

a) Parte I: formación

La primera parte corresponde a una etapa de formación básica en los temas que se han expuesto en la sección 4. Para realizar esta formación, se organizan un conjunto de seminarios, impartidos por los propios miembros del GrupoSI, más la participación de algunos “antiguos” colaboradores. La asistencia a estos seminarios es obligatoria para todos los estudiantes (tanto los de tipo (1) como los de tipo (2)). En estos seminarios se les proporcionará o se les indicará la bibliografía complementaria.

Al finalizar estos seminarios, se les realiza un *test* con un doble propósito. Por un lado, para motivar que los alumnos se vayan formando en los temas impartidos. Y por otro lado, para disponer de una primera “nota”. Este test incluye conocimientos básicos y lo suele aprobar cualquier estudiante que haya seguido los seminarios (ver sección 7).

b) Parte II: presentación de trabajos

Después de los seminarios, los colaboradores (1) realizan un trabajo que consiste en la búsqueda bibliográfica sobre un tema concreto, relacionado con alguno de los temas presentados en los Seminarios. Este tema se adecuará al nivel de cada estudiante (primer o segundo ciclo), y lo tendrán que presentar y exponer.

Por lo que respecta a los proyectistas (2), durante este período se concreta su proyecto y se le asigna un ponente (un miembro del GrupoSI afín al proyecto que va a realizar). A partir de aquí, será éste quien llevará personalmente al alumno.

c) Parte III: competición

Una vez presentados los trabajos, se propone a los colaboradores (1) un problema común que tienen que resolver individualmente o por grupos de dos. Cada grupo decide con qué técnica (GA, CBR o NN) se plantea resolverlo e implementarlo. La presentación del trabajo realizado consiste en: a) entregar una memoria donde se detalla y se justifica la resolución del problema realizada; y b) realizar una *competición*. Ésto se llevará a cabo sobre un mismo entorno desarrollado por el GrupSI (ver Parte II del artículo presente). “Ganará” el grupo que encuentre la mejor solución y obtenga un mejor rendimiento. Los parámetros concretos del problema no los conocerán hasta el mismo día de la competición. Así pues, deberán tener la “habilidad” de ajustar sus algoritmos *in situ* con tal de ser los mejores.

Cabe destacar que, de los miembros del GrupoSI se designa un responsable para cada nuevo curso, y se asignan “tutores/consultores” para cada área (GA, CS, GP, NN y CBR), a los cuáles los alumnos de

⁴⁹ Los alumnos que siguen trabajando con el GrupoSI en los años posteriores, realizan un trabajo más personalizado y ad hoc a su perfil.

los Seminarios pueden acudir en los horarios correspondientes. De la misma manera, los alumnos tienen un soporte telemático de los seminarios, donde pueden encontrar información general, documentación, notificaciones, calendario, etc.

6.- PLANIFICACIÓN GLOBAL

Una vez definidos los contenidos de los Seminarios y el planteamiento de los mismos, realizamos una planificación global de éstos. Concretamente, la tabla 3 muestra, como ejemplo, la planificación que se ha realizado para el curso académico 1999/2000, ya que en EALS los planes de estudio consisten en asignaturas anuales. Podemos destacar la particularidad que los cursos de las titulaciones de primer ciclo se evalúan en tres trimestres, mientras que las titulaciones correspondientes a un segundo ciclo se estructuran de forma cuatrimestral. Así pues, se han planificado los Seminarios respetando los períodos de exámenes de los distintos alumnos.

Tabla 3. Planificación global de los Seminarios del Grupo de Investigación en Sistemas Inteligentes para el curso 1999/2000.

Nº	Semana	Sesión (1 hora)	Sesión (1 hora)
1	18-22 octubre	Presentación / GA (I)	GA (II)
2	25-29 octubre	NN (I)	NN (II)
3	1-5 noviembre	NN (III)	GP
4	8-12 noviembre	CBR (I)	CBR (II)
5	15-19 noviembre	Herramienta GA	Herramienta NN
6	22-26 noviembre	CS	Día de trabajo personal
7	29-3 nov./dic.	Clase de dudas	Test
8	6-10 diciembre	Período de exámenes de primer ciclo	Período de trabajo personal de los de segundo ciclo
9	13-17 diciembre		
10	20-24 diciembre		
Fiestas de Navidad			
11	10-14 enero	Presentación de los trabajos de segundo ciclo	Período de trabajo personal de los de primer ciclo
12	17-21 enero		
13	24-28 enero	Período de exámenes de segundo ciclo	Período de trabajo personal de los de primer ciclo
14	31/4 enero/febrero		
15	7-11 febrero		
16	14-18 febrero	Presentación de los trabajos de primer ciclo	
17	21-25 febrero		
	marzo a junio	Colaboradores (1): implementación de un problema a resolver	Proyectistas (2): perfilan y definen su proyecto
18	Finales de junio	COMPETICIÓN	

7.- CRITERIOS DE EVALUACIÓN

Los criterios de evaluación que aquí presentamos, corresponden a los que se aplican a los alumnos que cursan una asignatura de libre elección dentro de los Seminarios (Intensificación y/o Especialización en Tecnología del Software). Concretamente, la nota final pretende ponderar los aspectos siguientes:

1. *Parte I: formación.* La asistencia a los seminarios, la participación en los seminarios, y la nota obtenida en el test. La nota del test es decisiva a efectos de decidir si el estudiante continua realizando normalmente los Seminarios. Consideramos que los estudiantes que lo suspenden no tienen ningún interés por el grupo. A este tipo de estudiantes se les proponen otro tipo de trabajos directamente para Setiembre.
2. *Parte II: presentación de los trabajos.* Evaluación del “trabajo-teórico” realizado, y evaluación de la presentación del mismo.
3. *Parte III: competición.* Evaluación de la memoria del “trabajo-práctico” realizado, y evaluación de la propia competición.

8.- CONCLUSIONES

Anteriormente al planteamiento de los Seminarios del GrupoSI (los cuáles se introdujeron por primera vez en el curso académico 1997/1998), los alumnos colaboraban directamente con algún tema concreto y con un profesor/a dado. Las desventajas que esto conllevaba eran que: 1) el alumno no adquiría una visión global de los distintos temas que se trataban dentro del GrupoSI y, al mismo tiempo, no tenía la sensación de estar integrado dentro de un grupo de personas con los mismos intereses; y 2) por parte del profesor/a, éste dedicaba mucho tiempo de su trabajo para un único alumno –o para un grupo muy reducido de personas-, casi equivalente al número de horas que se destinan a una asignatura de libre elección con los mismos créditos. Así pues, los Seminarios –aunque necesitan un trabajo continuo por nuestra parte- creemos que son positivos en distintos aspectos:

1. Los alumnos salen mejor preparados para una futura integración en el GrupoSI, ya que se les dan nociones básicas de diferentes temas en los cuáles estamos trabajando. Aunque su trabajo lo realicen sobre un tema concreto, el planteamiento de los seminarios les permite profundizar en el tema que más les interese.
2. Los alumnos sienten que pertenecen a un grupo de trabajo y esto, en general, los motiva más.
3. Los alumnos –según han comentado ellos mismos- encuentran una motivación complementaria. Una de las claves ha sido precisamente la *competición* que organizamos que, al fin y al cabo, sirve para realizar el mismo trabajo con cierta motivación para encontrar la mejor solución.
4. A los distintos miembros del GrupoSI nos permite organizar mejor nuestro tiempo y participar en los Seminarios con los temas que tenemos más afines.

Globalmente, aunque los Seminarios van evolucionando cada curso académico, creemos que éstos son positivos y constructivos para los alumnos que quieren definir su perfil en Sistemas Inteligentes como forma complementaria a las asignaturas de las titulaciones afines a nuestro grupo [6,8].

BIBLIOGRAFÍA

- [1] E. Bernadó, A. Mekaouche, J.M. Garrell, *A Study of a Genetic Classifier System Based on the Pittsburgh Approach on a Medical Domain*, IEA/AIE-99, 1999, Springer Verlag, 175-184
- [2] G. Calzada, J.M. Garrell, R. Barti, *ModelGP: A GP-based platform for predictive model construction*, Proceedings of SOCO'99, 1999
- [3] J.M. Garrell, E. Golobardes, E. Bernadó and X. Llorà, *Automatic Classification of Mammary Biopsy Images with Machine Learning Techniques*, Proceedings EIS'98, Volume 3: Artificial Intelligence, ICSS Academic Press, 1998, 411-418
- [4] J.M. Garrell, E. Golobardes, E. Bernadó and X. Llorà, *Automatic diagnosis with genetic algorithms and case-based reasoning*, Artificial Intelligence in Engineering, Elsevier Science Ltd. Vol. 13, No. 4, October 1999, 367-372, ISSN: 0954-1810
- [5] E. Golobardes, E. Pous, M. Román. *Introducció a les xarxes neuronals, presentació de les Backpropagation*, INPUT 10, Febrero 1996, 17-26
- [6] E. Golobardes, *¿Dónde están los límites de una introducción a la Inteligencia Artificial?* JENUI'97, Junio 1997, 293-300
- [7] E. Golobardes, J.M. Garrell, *Avaluació d'un sistema classificador basat en casos per a la diagnosi de biòpsies de teixit de glàndules mamàries*. Jornades d'Intel·ligència Artificial: Noves Tendències (JIA'97), ACIA No. 12, Tardor 1997, 41-49
- [8] E. Golobardes, J.M. Garrell, *¿Qué relación hay entre la asignatura de IA, la investigación y la aplicación de sus tecnologías?*, JENUI'98, Julio 1998, 347-354
- [9] X. Llorà, J.M. Garrell, *GENIFER: A Nearest Neighbour based Classifier System using GA*, Proceedings of GECCO'99, 1999, CA: Morgan Kaufmann, pp. 797
- [10] X. Llorà, E. Golobardes, M. Salamó, J. Martí, *Diagnosis of microcalcifications using Case-Based Reasoning and Genetic Algorithms*, Proceedings EIS'2000, To appear.
- [11] J. Martí, J. Español, E. Golobardes, J. Freixenet, R. Garcia, M. Salamó, *Classification of microcalcifications in digital mammograms using Case-Based Reasoning*, Proceedings of IWDM-2000, To appear.
- [12] M. Salamó, E. Golobardes, J.M. Garrell, J. Martí. *Clasificación de microcalcificaciones usando razonamiento basado en casos*, Proceedings CAEPIA'99/TTIA'99, 1999, ISBN: 931170-0-5