

Los observatorios virtuales astronómicos como nuevo campo de aplicación de la minería de datos

Luis Manuel Sarro

Dept.de Inteligencia Artificial

ETS Ingeniería Informática

Universidad Nacional de Educación a Distancia

c/ Juan del Rosal, 16

28040 Madrid

lsb@dia.uned.es

Enrique Solano

Observatorio Virtual Español

Laboratorio de Astrofísica Espacial

y Física Fundamental (LAEFF)

Estación de seguimiento de satélites de ESA

Villafranca del Castillo - Madrid

Enrique.Solano@esa.int

Resumen

En este artículo se presenta el campo de la Astrofísica como un área especialmente idónea para la aplicación de técnicas de Minería de Datos en la obtención de conocimiento. En primer lugar, se describen las entidades, propiedades y relaciones que aparecen en casi cualquier escenario del área. A continuación, se describen los Observatorios Virtuales, instituciones de reciente creación que han permitido la aplicación del corpus de conocimiento del área de Minería de Datos para la explotación científica de los archivos astronómicos. Finalmente se describen casos científicos prácticos que ilustran diversas tareas de Minería de Datos aplicadas a problemas astrofísicos.

1. Motivación

Así como en el área de la Medicina nos encontramos habitualmente con descripciones de problemas expresados en términos de pacientes, síntomas, enfermedades o tratamientos, en Astrofísica se maneja un lenguaje especializado cuyos rasgos más generales pretendemos esbozar aquí con el fin de hacer mínimamente comprensible nuestra exposición del campo como un área idónea para la evaluación de algoritmos y técnicas de descubrimiento de conocimiento. Todo ello sin pretender obvia-

mente que esta revisión sea completa o exhaustiva.

En Astrofísica, las entidades más habituales son las estrellas y las galaxias. Estas últimas son agregados de estrellas. Las estrellas aparecen en diferentes estadios evolutivos algunos de los cuales no son llamados propiamente estrellas: por ejemplo, una estrella en sus últimas fases de vida puede ser llamada nebulosa planetaria, enana blanca o púlsar. Las estrellas nacen a partir de una nube de gas y polvo que tampoco puede ser llamada propiamente estrella. Existen muy diferentes clases de estrellas dependiendo de sus características físicas (masa, temperatura, composición química...) y de su estado evolutivo (estrella en formación, estrellas de la secuencia principal, enanas blancas...). Vemos que la coordenada temporal desempeña un papel relevante en las descripciones de muchos escenarios astrofísicos.

Las galaxias son conglomerados de estrellas, de muy diferentes propiedades físicas y estados evolutivos. Las propiedades físicas y los estadios evolutivos de las galaxias se definen en términos estadísticos en función de los de las estrellas que las componen. Además, se incluye en la definición de galaxia el material interestelar que, como se mencionaba anteriormente, consiste principalmente de nubes de gas y polvo.

El espacio de parámetros en el que se definen los problemas astrofísicos se caracteriza

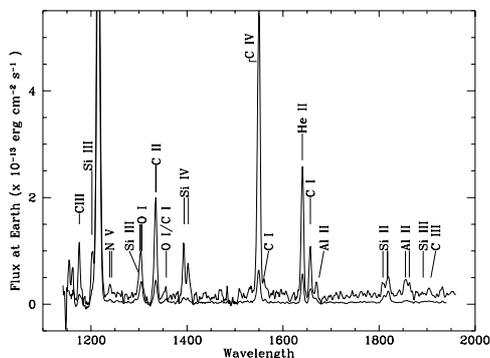


Figura 1: Espectro ultravioleta de la estrella II Pegasi.

por una dimensionalidad elevada y una heterogeneidad notable. En primer lugar, es importante hacer notar que, en general, de los objetos del Universo sólo disponemos información procedente de la luz que emiten. Hasta hace pocas décadas, los detectores sólo eran capaces de detectar luz en las longitudes a las que nuestros ojos también son sensibles (lo que comúnmente se denomina el rango óptico del espectro electromagnético). En las últimas décadas se han desarrollado detectores de luz infrarroja, ultravioleta, de rayos X y gamma, de ondas de radio... Con ello, hemos ampliado extraordinariamente el espacio de parámetros que describen los objetos astronómicos. Un elemento utilizado habitualmente en la clasificación de objetos astrofísicos es su espectro. Un espectro es un vector que contiene información sobre cuánta energía emite el objeto en una determinada longitud de onda (wavelength en inglés). La figura 1 muestra un espectro en el rango ultravioleta de una estrella activa.

En ella podemos observar cómo en determinadas longitudes de onda se emite mucha más energía que en otras. Por ejemplo, alrededor de 1550 Å se emite mucha energía, lo que los astrónomos interpretamos como indicio de la existencia de iones de carbono que han perdido tres electrones debido a las colisiones con otros átomos y electrones. Esa emisión prominente a 1550 Å es lo que llamamos una línea espectral. Los espectros se pueden emplear co-

mo atributos en Minería de Datos tanto en su versión sin procesar (un vector de flujos) como en versiones más procesadas en las que se emplean variables binarias que codifican la presencia/ausencia de una determinada línea espectral o variables discretas que codifican la cantidad de energía emitida en esa línea.

Por otro lado, los astrónomos han desarrollado técnicas para derivar parámetros físicos secundarios a partir de las observaciones de la luz de los diferentes objetos. Por ejemplo, se puede calcular la temperatura de la superficie de una estrella o del gas de una nube interestelar a partir de la luz que emite, o la composición química de cualquiera de esos u otros objetos (como ya hemos visto en el caso del carbono ionizado). También se puede inferir la distancia a la que se encuentra un objeto aunque la distancia sea uno de los parámetros secundarios más difíciles de inferir y que mayores errores relativos presenta. Todas ellas son susceptibles de ser utilizadas como atributos en Minería de Datos. De nuevo, partimos casi siempre de variables continuas que en determinadas circunstancias pueden ser aproximadas por variables discretas o binarias. En muchos casos además (aunque no en todos), se dispone de los errores asociados a tales medidas.

Las observaciones se pueden presentar, como hemos dicho, en forma de espectros, en los que tenemos información sobre la luz emitida en diferentes longitudes de onda, o en forma de magnitudes, en las que se suma toda la luz emitida por el objeto en un rango de longitudes de onda. Por supuesto, los primeros contienen una cantidad de información mucho mayor que las segundas pero el precio a pagar es una mayor dificultad técnica para su obtención. De cualquiera de ambos se puede tener una serie temporal. La figura 1 en realidad muestra dos espectros, uno de la estrella en calma y otro de la estrella durante un proceso conocido como fulguración durante el que las líneas de campo magnético se reconectan y liberan gran cantidad de energía. Otro ejemplo de la importancia de la coordenada temporal en la investigación astrofísica se ilustra en la figura 2 en la que se muestra un índice de actividad solar en función del tiempo. La actividad

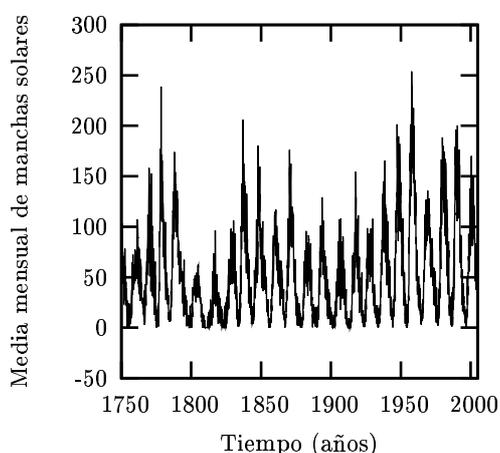


Figura 2: Serie temporal del número de manchas en la superficie del Sol.

solar, de origen magnético también, produce pérdidas importantes en la industria mundial por los efectos que tienen fulguraciones como la descrita en este mismo párrafo sobre las redes de comunicaciones y alta tensión. Asimismo, se ha demostrado ya la influencia de dicha actividad sobre el clima en la Tierra. Por ello, es importante predecir estas fulguraciones con tiempo suficiente como para prevenir sus efectos.

Así pues, podemos afirmar que en el dominio de la astrofísica tratamos con instancias que representan objetos del Universo, caracterizados por una serie de atributos entre los que cabe esperar variables continuas principalmente, pero discretizables en algunos casos, de alta dimensionalidad, asociadas a medidas directas del brillo del objeto, o derivadas de éstas. Las podemos encontrar en forma de variables, vectores, imágenes o cubos de imágenes. Se dan habitualmente casos de incompletitud en los atributos.

2. El observatorio virtual: concepto y objetivos.

El objetivo final de toda misión científica es la obtención de datos de gran calidad. Para

rentabilizar de forma eficiente la utilización de dichos datos, es necesario desarrollar sistemas que garanticen su perdurabilidad y permitan un fácil acceso a los mismos. La Astrofísica no ha sido ajena a esta necesidad y todas las misiones espaciales, así como la mayoría de los grandes observatorios terrestres, han desarrollado en los últimos años sistemas eficientes de archivo y acceso a los datos científicos que facilitan al usuario la identificación de información de interés para su posterior análisis. Es importante, llegados a este punto, hacer notar la diferencia entre el simple almacenamiento seguro de los datos, que tendría un carácter pasivo, y un verdadero archivo científico, el cual requiere una catalogación precisa de los contenidos así como el desarrollo e implementación de herramientas de búsqueda, presentación y transferencia de los datos.

Es indudable que, hoy en día, los archivos de datos constituyen una pieza fundamental en la investigación Astrofísica tal y como refleja el uso intensivo que, de estos recursos, hace la comunidad científica. La existencia de los archivos permite abordar líneas de investigación que, de otra manera, serían inviables como, por ejemplo, estudios de variabilidad temporal. En este sentido es de destacar el ejemplo de INES, el sistema de acceso, archivo y distribución de los datos del satélite IUE, en el que el número de peticiones de datos al archivo en los últimos tres años ha superado el número de observaciones realizadas por el satélite en sus 18 años de misión, cifrado en 104000.

Hasta ahora, la comunidad científica del área sólo disponía de información en el rango visible del espectro electromagnético de forma que se ha construido *de facto* una ontología basada en aspectos parciales de las entidades objeto de estudio. En particular, las clases se han definido a partir exclusivamente de las características fenomenológicas en el rango visible. Con el advenimiento de las misiones espaciales en otros rangos del espectro (el ultravioleta, el infrarrojo, los rayos X y gamma, etc) se ha completado un cuadro que antes sólo podíamos contemplar parcialmente. Sin embargo, la ingente acumulación de datos sobre

los millones de objetos observados bien desde instalaciones terrestres, bien desde el espacio se almacenaba en el mejor de los casos de forma totalmente heterogénea. No existía un modelo de datos estándar aceptado por la comunidad. Como resultado, cada archivo científico empleaba formatos distintos, nomenclatura diferente (de forma que un mismo objeto podía ser llamado de hasta veinte formas distintas), esquemas clasificatorios variados (de forma que un mismo objeto puede ser clasificado de muchas maneras distintas)... En particular, no existía un modelo de datos unificado capaz de integrar espectros en diferentes longitudes de onda, imágenes bidimensionales, cubos de datos, magnitudes fotométricas, series temporales y parámetros físicos secundarios, por citar sólo algunos ejemplos de atributos de esos objetos. Sólo con la llegada de los observatorios virtuales se ha conseguido aunar toda esa información dispersa y presentarla de forma consistente e interoperable. Todo ello permite en la actualidad explorar problemas científicos cuya formulación era imposible antes de lo que se conoce como astronomía multirrango (en alusión a la posibilidad de combinar las observaciones realizadas en distintos rangos del espectro electromagnético). Sin embargo, su resolución pasa inevitablemente por la capacidad de manejar de forma eficiente un conjunto estadísticamente significativo (es decir, con un número grande de instancias) de objetos cuyos parámetros proceden de diferentes fuentes. La información ya disponible en los distintos archivos astronómicos junto con las grandes campañas de observación en curso o previstas para los próximos años permitirán tener una cobertura de grandes zonas del cielo en la mayor parte del espectro electromagnético. Es, por tanto, indudable que una coordinación eficiente entre los distintos centros de datos impulsaría de una manera sustancial la producción científica, rentabilizando al máximo el elevado coste que implican dichos proyectos.

Un problema adicional a la falta de interoperatividad lo constituía la gestión del enorme volumen de datos producidos por proyectos de observación global del cielo. Proyectos como

SDSS y 2MASS, con millones de objetos observados, son una buena muestra de esto. El estudio de tales cantidades de datos no puede realizarse siguiendo los procedimientos tradicionales de transferencia de información desde el centro de datos para su posterior análisis local. Así por ejemplo, suponiendo una más que aceptable capacidad de transmisión de 1 Mb/s, se necesitarían dos semanas para transferir el catálogo del DR2 (Data Release 2, uno de los subconjuntos de datos ya públicos del SDSS) y dos meses para la transferencia de las imágenes las cuales ocuparían unos 1000 DVDs. La situación será incluso peor al final del proyecto ya que la cantidad de datos será aproximadamente el triple de la producida para el DR2. Aún más, el problema no solamente se ciñe al contexto de la transferencia de datos sino que afecta igualmente al propio análisis de los mismos ya que la mayoría de las herramientas utilizadas comúnmente por la comunidad astronómica no son escalables a las cantidades de datos generadas por estos proyectos.

El Observatorio Virtual nace con el objetivo de solventar los problemas que para la astronomía multirrango crea la falta de interoperatividad existente entre los centros de datos astronómicos y de proporcionar a la comunidad astronómica internacional los recursos necesarios para poder realizar una explotación eficiente de la cantidad ingente de datos que albergan los distintos archivos. El proyecto surgió hacia el año 2000 como resultado de iniciativas tanto nacionales (NVO, AstroGrid) como supranacionales (AVO). En Junio de 2002 surge la International Virtual Observatory Alliance (IVOA) con el propósito de ser un marco común en donde se coordinen e integren todas estas iniciativas. En la actualidad se encuentra formada por 15 proyectos de índole tanto nacional como supranacional. Dentro de IVOA se han definido las líneas de actuación del Observatorio Virtual hasta 2005, en donde destacan las demostraciones científicas que, desde 2003, se vienen realizando de forma anual con el doble objetivo de dar a conocer el proyecto a la vez que demostrar el enorme impacto que puede tener en la inves-

tigación astronómica. Información exhaustiva sobre la última de estas demostraciones se puede encontrar en <http://www.eurovo.org/twiki/bin/view/Avo/SwgMeeting06>. Aún siendo un proyecto de muy reciente creación, ya ha superado la fase de demostración tecnológica y constituye una herramienta de enorme utilidad en la investigación astronómica. Un ejemplo lo constituye el descubrimiento de agujeros negros supermasivos a partir de datos de GOODS (Great Observatories Origins Deep Survey, Padovani et al. 2004, 424, 545)

En el contexto del Observatorio Virtual, un centro de datos astronómicos no puede ser simplemente un centro que se limite a proporcionar datos en distintas etapas de su proceso de reducción sino que debe implementar toda una serie de utilidades que faciliten al máximo la labor de selección y posterior análisis de los datos. Este tipo de utilidades representan un valor añadido fundamental para el sistema y tienen un doble efecto: por un lado permite una utilización general por parte de la comunidad científica y, por otro lado, evita al usuario tener que dedicar tiempo al desarrollo de las mismas, máxime cuando, en la mayoría de los casos, dicho desarrollo requiere un conocimiento tanto científico como técnico no trivial.

Esta nueva concepción de Centro de Datos Científicos se encuentra plenamente justificada por la enorme cantidad de datos que los distintos proyectos astronómicos de observación de grandes regiones del cielo, en curso o previstos para un futuro cercano (SDSS, 2MASS, VISTA...), van a proporcionar durante su vida operacional. La ingente cantidad de información que los archivos de estos proyectos van a almacenar supera ampliamente las posibilidades de los investigadores para interpretarlos y analizarlos siguiendo el modelo tradicional según el cual el usuario accede al centro de datos, selecciona aquéllos que son de su interés y los transfiere a través de Internet a su estación de trabajo en donde realiza el análisis de los mismos utilizando una serie de herramientas estándar. En el nuevo escenario, el análisis de los datos se realiza remo-

tamente en el Centro de Datos y solamente se transfieren aquéllos que, una vez analizados, son de verdadero interés para el usuario. Este planteamiento presenta la gran ventaja de que el usuario no necesita transferir grandes volúmenes de datos a su estación de trabajo con los inconvenientes y limitaciones que este planteamiento tiene tanto desde el punto de vista de la transferencia de datos como desde el conocimiento necesario para poder aplicar las herramientas de análisis adecuadas. VOSTAT, un servicio Web que permite la utilización de herramientas estadísticas complejas sobre grandes conjuntos de datos es un magnífico ejemplo del concepto de herramienta que se describe en esta sección.

La existencia de métodos automáticos implementados como servicios Web tiene además otras ventajas adicionales como la homogeneidad, y la inmunidad contra errores asociados a la existencia de un grupo heterogéneo de investigadores. A nuestro entender, el Observatorio Virtual, al garantizar la interoperatividad entre diferentes conjuntos de datos, constituye el marco ideal en donde desarrollar este tipo de técnicas.

3. Tareas de descubrimiento de conocimiento en astrofísica.

Un proyecto como el Observatorio Virtual donde se pretende relacionar cantidades ingentes de datos en distintas ventanas del espectro electromagnético, es el marco ideal para desarrollar herramientas de Minería de Datos. Es precisamente mediante la utilización de estas técnicas de donde se espera obtener los principales descubrimientos científicos.

Así pues, el empleo de herramientas de Minería de Datos para la explotación científica de observatorios virtuales tiene como objetivo posibilitar el descubrimiento científico en problemas de alta dimensionalidad o con un elevadísimo número de instancias. Hasta el momento presente se han identificado varias tareas de especial interés en el ámbito astrofísico. Son tareas clásicas en la disciplina de la Minería de Datos como la búsqueda de reglas de asociación, la clasificación, la regresión, la

predicción de series temporales y, finalmente, el análisis de agrupamientos o clustering y la detección de objetos exóticos. Como ejemplos de aplicación en los campos mencionados anteriormente podemos citar los siguientes:

1. Agrupación de datos no supervisada. Se trata quizás de la tarea más novedosa dentro del campo de la astrofísica. Se ha abordado desde una perspectiva conexionista mediante mapas autoorganizados ([1],[2]). Dada la proliferación de lo que se conoce en inglés como *surveys* (exploraciones sistemáticas de todos los objetos de una región del cielo o de todo el cielo hasta un cierto límite de brillo), se hacen cada vez más necesarios métodos de análisis que nos permitan segmentar bases de datos de $10^8 - 10^9$ objetos en grupos correlacionados. Los objetos exóticos (*outliers*) son de especial interés en astrofísica en la medida en que abren escenarios físicos no estándar. Más aún, resulta de especial interés un tipo de consulta basado en similitud, consistente en pedir a un observatorio virtual los objetos próximos en el espacio de parámetros a uno dado.
2. Obtención de leyes correlacionales que permitan obtener parámetros físicos complejos a partir de observaciones fácilmente accesibles. Un ejemplo clave de este tipo de leyes es la relación periodo-luminosidad en estrellas pulsantes cefeidas que nos permite calcular la distancia a una de estas estrellas una vez conocido el periodo de la pulsación. El periodo es muy fácil de obtener simplemente midiendo el brillo de la estrella durante varios días. Sin embargo, la distancia es la magnitud más difícil de obtener en astrofísica. Dentro de este tipo de estudios o relacionado de alguna manera con él, existe un tipo de tarea muy importante consistente en seleccionar el conjunto de atributos óptimo para la identificación de un tipo de objetos. Esta tarea puede considerarse un híbrido entre la clasificación, la búsqueda de correlaciones o simplemente el filtrado selectivo de atributos y se emplea con diversos fines, entre otros, para ampliar las muestras de clases de objetos exóticos para poder elaborar estudios estadísticos sobre sus propiedades.
3. Clasificación automática de objetos basada en datos (esto es, objetiva, cuantificable y reproducible) y no en etiquetas. Esta última ha sido hasta ahora la única accesible para los usuarios de archivos astronómicos. El astrónomo que solicitó originalmente las observaciones archivadas las clasifica de forma arbitraria, de acuerdo a su aproximación parcial al objeto observado y basada únicamente en los datos a su disposición. Se han empleado casi exclusivamente redes neuronales (perceptrones multicapa obtenidos mediante retropropagación del error) para clasificar morfológicamente galaxias a partir de imágenes bidimensionales ([3], [4], [5] o [6]); para distinguir estrellas de galaxias ([7], [8], [9] o [10]); para clasificar estrellas ([11], [12], [13], [14] or [15]) o galaxias ([16]) a partir de sus espectros; o, finalmente, para clasificar series temporales ([17]).
4. Regresión de parámetros físicos a partir de observaciones. Un ejemplo paradigmático consiste en predecir las propiedades físicas de la atmósfera de una estrella a partir de la radiación que emite (su espectro). Los parámetros predichos pueden ser temperaturas, gravedades o metalicidades [18] o incluso vectores de campo magnético, velocidades proyectadas y características mucho más complejas, siempre ligadas de forma no lineal con los parámetros observables de la estrella([19]). Casi siempre, este tipo de tareas se aborda desde la perspectiva conexionista mediante perceptrones multicapa entrenados con retropropagación del error.
5. Predicción de series temporales, en particular relativa a la actividad solar ([20] es una revisión interesante de la bibliografía al respecto).

4. Conclusión

El Plan Nacional de Astronomía y Astrofísica, dentro de sus líneas de investigación prioritarias, recoge la necesidad de la explotación científico-tecnológica de los recursos astronómicos, en particular los relacionados con los centros de archivo y observatorios virtuales. Esta línea de actuación se encuentra en consonancia con la recomendadas por la Unión Europea en su Libro Blanco 'Space: a new European frontier for an expanding Union. An action plan for implementing the European Space policy' y que se transcribe a continuación:

'Space science data represent a significant long-term investment. They have to remain affordable and accessible over a period extending well beyond the nominal lifetime of the missions. There is a need to develop and sustain new infrastructures - thematic centres and networks- for long term archiving and distribution of European space data. This infrastructure development should allow for integration of space and non-space data archives, a key issue for multidisciplinary work. Recommended actions: Progressively increase ESA and national funding for space research... in particular with the objectives of ... developing the supporting infrastructure for data acquisition, long-term archiving and dissemination.'

En España contamos ya con un Observatorio Virtual que proporciona acceso a la comunidad científica internacional, a los datos de varias misiones espaciales y que contará en pocos años con el archivo de las observaciones realizadas por uno de los pocos telescopios de más de 10 metros de diámetro del mundo (GRANTECAN en Canarias). En la actualidad, el SVO (siglas en inglés del Observatorio Virtual Español) está realizando desarrollos pioneros de herramientas de Minería de Datos en este ámbito, basados en distintos tipos de clasificadores (bayesianos, máquinas de vectores soporte, redes neuronales bayesianas...) y sistemas de ayuda a la decisión en el campo de la interpretación de observaciones mediante modelos teóricos. El objetivo final es dotar al SVO de un centro de análisis científico

que facilite la utilización de tecnología de vanguardia en el área de la Minería de Datos para la explotación de los archivos astronómicos. En ese sentido, solicitamos la colaboración de los grupos españoles del área de la Minería de Datos para el desarrollo de aplicaciones en el ámbito de la Astrofísica que consoliden una posición de liderazgo del SVO dentro de la Alianza Internacional de Observatorios Virtuales y del campo de la explotación de bases de datos científicas.

Referencias

- [1] Brett, D.R., West, R.G., Wheatley, P.J.: The automated classification of astronomical light curves using Kohonen self-organizing maps. *MNRAS***353** (2004) 369–376
- [2] Rajaniemi, H.J., Mähönen, P.: Classifying Gamma-Ray Bursts using Self-organizing Maps. *ApJ***566** (2002) 202–209
- [3] Odewahn, S.C., Cohen, S.H., Windhorst, R.A., Philip, N.S.: Automated Galaxy Morphology: A Fourier Approach. *ApJ***568** (2002) 539–557
- [4] Goderya, S.N., Lolling, S.M.: Morphological Classification of Galaxies using Computer Vision and Artificial Neural Networks: A Computational Scheme. *Ap&SS***279** (2002) 377–387
- [5] Somanah, R., Rughooputh, S.D.D.V., Rughooputh, H.C.S.: Identification and classification of galaxies using a biologically-inspired neural network. *Ap&SS***282** (2002) 161–169
- [6] Bazell, D., Aha, D.W.: Ensembles of Classifiers for Morphological Galaxy Classification. *ApJ***548** (2001) 219–223
- [7] Philip, N.S., Wadadekar, Y., Kembhavi, A., Joseph, K.B.: A difference boosting neural network for automated star-galaxy classification. *A&A***385** (2002) 1119–1126

- [8] Cortiglioni, F., Mähönen, P., Hakala, P., Frantti, T.: Automated Star-Galaxy Discrimination for Large Surveys. *ApJ***556** (2001) 937–943
- [9] Andreon, S., Gargiulo, G., Longo, G., Tagliaferri, R., Capuano, N.: Wide field imaging - I. Applications of neural networks to object detection and star/galaxy classification. *MNRAS***319** (2000) 700–716
- [10] Mähönen, P., Frantti, T.: Fuzzy Classifier for Star-Galaxy Separation. *ApJ***541** (2000) 261–263
- [11] Snider, S., Allende Prieto, C., von Hippel, T., Beers, T.C., Sneden, C., Qu, Y., Rossi, S.: Three-dimensional Spectral Classification of Low-Metallicity Stars Using Artificial Neural Networks. *ApJ***562** (2001) 528–548
- [12] Allende Prieto, C., Rebolo, R., López, R.J.G., Serra-Ricart, M., Beers, T.C., Rossi, S., Bonifacio, P., Molaro, P.: The INT Search for Metal-Poor Stars: Spectroscopic Observations and Classification via Artificial Neural Networks. *AJ***120** (2000) 1516–1531
- [13] Weaver, W.B.: Spectral Classification of Unresolved Binary Stars with Artificial Neural Networks. *ApJ***541** (2000) 298–305
- [14] Bailer-Jones, C.A.L., Irwin, M., von Hippel, T.: Automated classification of stellar spectra - II. Two-dimensional classification with neural networks and principal components analysis. *MNRAS***298** (1998) 361–377
- [15] Vieira, E.F., Ponz, J.D.: Automated classification of IUE low-dispersion spectra. I. Normal stars. *A&AS***111** (1995) 393–+
- [16] Folkes, S.R., Lahav, O., Maddox, S.J.: An artificial neural network approach to the classification of galaxy spectra. *MNRAS***283** (1996) 651–665
- [17] Balastegui, A., Ruiz-Lapuente, P., Canal, R.: Reclassification of gamma-ray bursts. *MNRAS***328** (2001) 283–290
- [18] Bailer-Jones, C.A.L.: Stellar parameters from very low resolution spectra and medium band filters. T_{eff} , $\log g$ and $[M/H]$ using neural networks. *A&A***357** (2000) 197–205
- [19] Carroll, T.A., Staude, J.: The inversion of Stokes profiles with artificial neural networks. *A&A***378** (2001) 316–326
- [20] Conway, A.J.: Time series, neural networks and the future of the Sun. *New Astronomy Review* **42** (1998) 343–394