Automatic learning of structural knowledge from geographic information for updating land cover maps

Meriam Bayoudh, Emmanuel Roux, Richard Nock, G. Richard

Automatic Learning of Structural Knowledge from Geographic Information for updating Land Cover Maps

Meriam Bayoudh[1,2]
Emmanuel Roux [1]
Richard Nock[2,3]
Gilles Richard[4]

[1] ESPACE-DEV, UMR228 IRD/UMII/UR/UAG, Institut de Recherche pour le Développement, Cayenne, French Guiana, France
meriam.bayoudh@ird.fr ; emmanuel.roux@ird.fr
[2] Université des Antilles et de la Guyane, France
[3] CEREGMIA, Schoelcher, Martinique, France
Richard.Nock@martinique.univ-ag.fr
[4] British Institute of Technology and E-commerce, London, United Kingdom
grichard@bite.ac.uk

Abstract. The number of satellites and remote sensing sensors devoted to earth observation becomes increasingly high, providing more and more data and especially images. In the same time the access to such a data and to the tools to process them has been considerably improved. In the presence of such data flow - and regarding the necessity to follow up and predict environmental and societal changes in highly dynamic socio-environmental contexts - we need automatic image interpretation methods. This could be accomplished by exploring some strengths of artificial intelligence. Our main idea consists in inducing classification rules that explicitly take into account structural knowledge, using Aleph, an Inductive Logic Programming (ILP) system. We applied our proposed methodology to three land cover/use maps of the French Guiana littoral. One hundred and forty six classification rules were induced for the 39 land-cover classes of the maps. These rules are expressed in first order logic language which make them intelligible and interpretable by non-experts. A ten-fold cross validation gave average values for classification accuracy, specificity and sensibility equal to, respectively, 98.82 %, 99.65% and 70%. The proposed methodology could be valuably exploited to automatically classify new objects and/or help operators using object-based classification procedures.

Keywords: Inductive Logic Programming (ILP), Machine learning, Remote sensing, Geographic Information System (GIS), Land cover/use maps

1. Introduction

Geographic information have always been a key element for decision making in public policy, especially for planning. However, such information and the tools for represent, analyse and interpret them, particularly Geographic Information Systems (GIS), have been becoming more and more popular for the last two decades. Among the geographic information sources, remote sensing data have been taking an increasing part, because of the multiplication of the earth observation satellites and sensors and the evolutions in data distribution policies, more and more countries and/or organizations distributing remotely sensed data for free. Such a data quantity makes the data processing and interpretation a new challenge for engineers and researchers. Therefore, the latter can not keep on applying classical procedures but need new approaches in order to, notably, automatically updating land cover/use maps that provide useful information to decision makers. Such automatic procedures have to be based on existing knowledge, coming not only from remote sensing and image processing experts but also from final users. Some works aim at formally represent and exploit these expert knowledge for automatic image classification and interpretation : Desachy already proposed a fuzzy expert system in 1991, called ICARE, (DESACHY, 1991). More recent studies are directed rather towards ontologies. For instance, in (HUDELOT ; ATIF ; BLOCH, 2008) an ontology of spatial relations is proposed in order to guide image interpretation, then this ontology is enriched by fuzzy representations of concepts. In addition, (DURAND et al., 2007) propose an approach based on ontology for object recognition in remote sensing images.

A complementary approach to the expert based one is to extract knowledge from data. However, The great majority of the methods devoted to satellite image supervised classification consider only the pixel information within the image regions belonging to the same class in order to learn class signatures. Structural aspects within the pixel neighbourhood are essentially taken into account by computing textural indexes within the same regions. Very few methods tend to learn structural symbolic knowledge at a higher semantic level, like the one constituted by regions that can be identified in the image, by means of expert intervention, the application of segmentation or of non-supervised classification procedures. Paradoxically, some software propose to the operator to integrate such high level symbolic knowledge within the classification process, within a methodological framework called *object-oriented classification*. With the huge number of possible combinations offered by such software, the non-expert user feels somewhat lost and tends to formulate too specific and not reproducible rules. To our knowledge, no learning procedure, within such software environments, is proposed to help such novice users defining more general and efficient rules exploiting structural aspects. However, some studies already propose to learn structural knowledge from existing maps, for different purposes. In (MALERBA et al., 2003), the authors propose to help interpreting topographic maps. Their system, called INGENS, integrates machine learning tools and GIS functionalities. GIS classical functionalities are used to extract relevant concepts and features from spatial database, and the integrated inductive system permits to find rules in order to automatically recognize complex geographical contexts (like "fluvial landscapes") defined by the presence of specific geographical objects and their spatial arrangement. In (VAZ ; FERREIRA ; LOPES, 2007), the authors use Inductive Logic system called APRIL (FONSECA ; SILVA ; CAMACHO, 2006) to learn classification rules, from, on one hand, a detailed map provided by botanists and, on the other hand, Corine Land Cover (CLC) maps of the same zone. Such rules are intended to automatically disaggregate CLC map information that are judged too generic within the application framework. An application of the inductive learning of structural features from maps are devoted to the prediction of particular events that partially depend on landscape characteristics : Vaz *and al.* (VAZ ; COSTA ; FERREIRA, 2010) propose a system in order to predict wildfires, from information on past fire events and from compositional and structural features of the land use.

In this context, our work consists in implementing machine learning methods for structural and symbolic knowledge extraction from land use/cover maps and from various complementary geographic information layers. We define structural knowledge, on the one hand, as the knowledge that concern intrinsic structures of the land cover/use classes (e.g : ”The class C contains small objects”), and on the other hand, as the knowledge that concern relations, in space and/or time, between "objects" (e.g : Objects of the class $C1$ are always adjacent to the class $C2$). We chose Inductive Logic Programming (ILP) to implement the machine learning thanks to clearness of the language used and the intelligibility of its resulting hypothesis. Inductive Logic Programming (MUGGLETON, 1991) can model complex problems and was used in several domains such as : Chemistry (BLOCKEEL et al., 2004), biology, physics, medicine (LUU et al., 2012; FROMONT ; CORDIER ; QUINIOU, 2005), ecology, bioinformatics (SANTOS et al., 2012; LAVRAC ; DZEROSKI, 1994; SRINIVASAN et al., 1996). It was applied for spatial data mining (CHELGHOUM et al., 2006), for Chess (GOODACRE, 1996) and to test quality of river water (CORDIER, 2005).

We applied the proposed methodology to the learning of classification rules for the updating of the land cover/use in French Guiana.

Our paper is organized as follows : Section 2 introduces the general methodology by defining ILP basis, presenting the used dataset and explaining how the extracted information are coded. Section 3 presents the results and is followed by a discussion in section 4 and a general conclusion in section 5.

2. Methodology

2.1. Inductive Logic Programming

Inductive Logic Programming (ILP) (MUGGLETON, 1991) is a search fields that combines machine learning and logic programming. It is a technique for learning a general theory H from a background knowledge B and examples E within a framework provided by clausal logic. More formally (LAVRAC ; DZEROSKI, 1994), ILP is defined as follows :

Given :

- Background knowledge B expressed under Horn clauses, describing a set of knowledge and constraints that concerns the target concept, i.e., in our case, the membership to a given land cover/use class ;
- A set of examples E, divided into two subsets : E^+ and E^- that represent the sets of, respectively, positive and negative examples ;
- A description language L.

Find a "theory" H expressed under logic program using the description language L and that covers positive examples E^+ and does not cover (or in a controlled quantity) the negative examples E^-.

We chose the ILP engine Aleph (SRINIVASAN, 2007). It is an open source ILP system written in Prolog, using top-down search and based on inverse entailment (MUGGLETON, 1995).

We set the minimum *accuracy* of the candidate clauses to 0.7, *accuracy* being defined by $p/(p+n)$ where p and n are respectively the number of positive examples and the number of negative examples covered by the clause, and the maximum clause length to 6 literals.

2.2. Dataset

We applied the proposed methodology to the follow up of French Guiana littoral (Figure 1). Such territory is subjected to intense anthropogenic and natural dynamics (ANTHONY et al.,

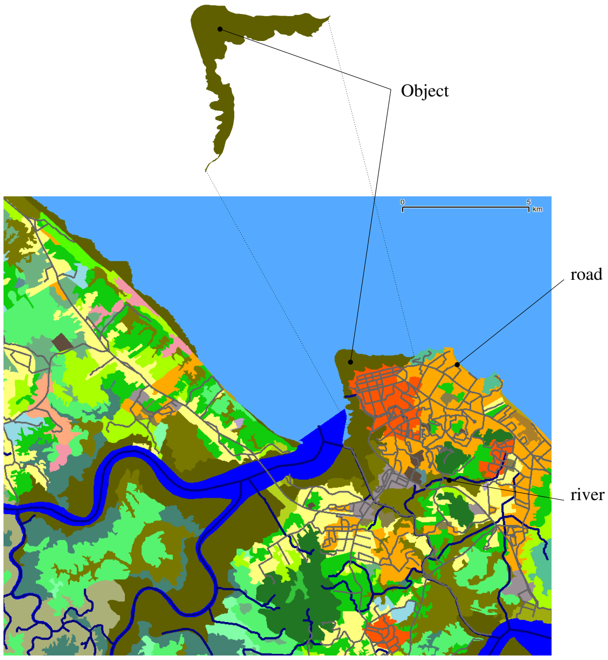
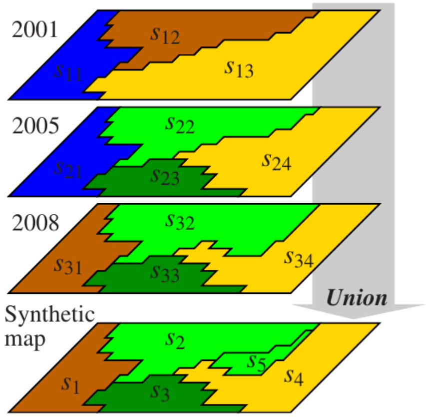


FIGURE 1 – Subset of the land cover/use maps and of the complementary geographic information layers used in this article. Sources : French National Office of Forests (Office National des Forêts : ONF), French National Geographic Institute (Institut Géographique National : IGN) ; French ministry in charge of the environment ; Regional Direction of the Environment (DIREN) of French Guiana ; French National Agency for Water and Aquatic Environments (ONEMA). See text for details.

2010). We benefited from a series of three land cover/use maps of the years 2001, 2005 and 2008. The classification nomenclature is based on the CORINE Land Cover (CLC) European nomenclature, adapted to the Amazonian context by the addition of fifteen classes, nine of them corresponding to different types of forests. It consists of three nested levels which the most detailed (level III) is compounded of 39 classes. The maps have been produced by the French National Office of Forests (Office National des Forêts : ONF) by photo-interpretation of the BD-Ortho® aerial photographs database of the French National Geographic Institute (Institut Géographique National : IGN) for the years 2001 and 2005. Air photographs have a 50-cm spatial resolution. The land cover/use map for 2008 was obtained by updating the anterior map using 2.5-meter spatial resolution satellite images acquired by SPOT 5 satellite and obtained through the SEAS-Guyane [1] project.

Two complementary geographic information layers were used (see Figure 1) : the road network, provided by the BD-Cartho® database of the IGN, and the river network provided by the BD-Carthage® database of the French ministry in charge of the environment and of the IGN, produced in 2009 for French Guiana by the Regional Direction of the Environment (DIREN) of French Guiana and the French National Agency for Water and Aquatic Environments (ONEMA).

1. https ://www.seas-guyane.org



Synthetic map attribute table

Object	Class (2001)	Class (2005)	Class (2008)[1]
s_1	blue	blue	brown
s_2	brown	green1	green1
s_3	orange	green2	green2
s_4	orange	orange	orange
s_5	orange	orange	green1

[1] Classes the membership of which is to be learned

FIGURE 2 – Illustrative example explaining the definition of the synthetic map which combines the three initial map information.

2.3. Geographic information extraction and coding

2.3.1. Data preprocessing / Definition of the map objects

First of all, we completed the initial land cover/use types by adding three more classes : *Ocean*, *River* and *Unknown*. The first two classes were added as they contribute significantly to the structuring of the environment on the French Guiana territory. The *Unknown* class was defined to explicitly take into account the fact that, for some areas, information were not available in 2001 and/or 2005. However, we did not induced rules in order to predict *Ocean* and *Unknown* class memberships. Eventually, the class *Paddy field* was not considered as it was under-represented in the maps.

At the next step, we produced a synthetic map by merging the information contained in the three initial maps, by means of the "union" GIS operator. The elementary geographical entities of the resulting map are referred to as *objects* in the following and contribute to define the *examples* in the ILP process. Objects constituting such a synthetic map belong to only one class at each date, as schematically shown in Fig. 2.

2.3.2. Information coding

Target predicates (i.e. concepts to be learned) were defined as the land cover/use classes to which the objects of the synthetic map belong in 2008.

Objects were described thanks to predicates characterizing intrinsic (class, area, fractal dimension, compacity, perimeter) and relational features (adjacency, inclusion, relative positions in latitudinal and longitudinal directions) :

- classYY($O1$,class_type) : indicates the land cover class of the object $O1$ at the year YY.
- adjacent($O1$,$O2$) : indicates that objects $O1$ and $O2$ are adjacent, *i.e.* that they share, at least partially, their boundary ;
- included($O1$,$O2$) : means that object $O2$ is included in object $O1$;
- contain($O1$,X), with $X \in \{road, river\}$: indicates the presence of a road or river in object $O1$;
- lat($O1$,y) and long($O1$,x) : y is a real and positive number corresponding to, respectively, the latitude and the longitude of the centroid of object $O1$.

However, absolute positions of the objects were not expected in induced rules. Relative positions of the objects in the latitudinal direction were coded as follows (the principle was the same for the longitudinal direction, by defining the predicates *east* and *west*) :

– north($O1$,$O2$) :- lat($O1$,A),lat($O2$,B),A$>$ B : indicates that object $O1$ is located at the north of object $O2$;
– south($O1$,$O2$) :- lat($O1$,A),lat($O2$,B),A$\leq$ B : indicates that object $O1$ is located at the south of object $O2$;

For each numeric descriptive attribute (area, fractal dimension, compacity, perimeter), object value was re-coded by indicating whether it is lower or equal (le) or greater (g) than the 10th, 20th, ... and 90th percentiles of the data empirical distribution. For instance, the area of an object $O1$ was re-coded by means of the following predicates :

– area($O1$,leP_i) :-area($O1$,Y), $Y \leq P_i$.
– area($O1$,gP_i) :-area($O1$,Y), $Y > P_i$.

P_i being the i^{th} percentile of the empirical distribution of the area data, and $i \in \{10, 20, ..., 90\}$.

All object features were extracted using the free and open source GRASS Geographic Information System (GRASS_DEVELOPMENT_TEAM, 1999-2012). Once the information is extracted and coded according the above method, the one vs. rest approach was applied to induce the rules, providing one classifier per land cover/use class.

3. Results

3.1. Set of induced rules

We obtained a total of 146 rules. The number of positive examples covered by a rule vary from 2 to 692 examples, while negative examples covered by a rule vary from 0 to 101 examples. Classification rule premisses are composed of 1 to 3 literals.

Below we show some induced rules with, in brackets, the numbers of positive (*Pos cover*) and negative (*Neg cover*) examples covered by the rule and the total number of positive examples for the considered target predicate (*Total pos. ex.*)

(1) `(Pos cover = 7 Neg cover = 3 Total pos. ex. = 1191)`
`class08(A, Isolated build) :- class05(A, Flooded forest or swamp), compacite(A,le1.21), perimeter(A,le 552.62).`
(2) `(Pos cover = 692 Neg cover = 36 Total pos. ex. = 1191)`
`class08(A,Isolated build) :- class05(A,Isolated build).`
(3) `(Pos cover = 198 Neg cover = 65 Total pos. ex. = 552)`
`class08(A,Multidisciplinary habitat) :- adjacent(A,B), class05(B,Multidisciplinary habitat), area(A, le 165566.67).`
(4) `(Pos cover = 2 Neg cover = 0 Total pos. ex. = 84)`
`class08(A,Road network) :- class01(A,Degraded forest), class05(A, Road network), compacity(A,g2.10).`
(5) `(Pos cover = 3 Neg cover = 1 Total pos. ex. = 97)`
`class08(A,Construction site) :- class05(A,Construction site), area(A,g10830.61), perimeter(A, le 780.61).`
(6) `(Pos cover = 3 Neg cover = 0 Total pos. ex. = 194)`
`class08(A,High forest) :- contains(A,Road), class05(A,High forest), area(A,le1111.13).`

To illustrate how we interpret ILP rules, consider the rule (3) : it covers 198 positive examples over 552 and only 65 negative examples over 6615. It states that an object A belongs to the class *Multidisciplinary habitat* in 2008 if A is adjacent to another object B that belonged to the same class in 2005 and if A has an area lower or equal to 165566.67 square meters.

TABLE 1 – Ten-fold cross validation results for land cover/use classes associated with "low" sensibility values (lower than 50%)

Class (level I)	Class (level II)	Class (level III)		Precision	Sensibility	Specificity	Total number of positive examples
Non-agricultural artificial territories	Mines, rubbish dump construction sites	Rubbish dump		99.79	25	99.96	15
		Construction sites		98.81	26.67	99.8	97
Agricultural territories	Heterogeneous agricultural areas	Fragmented/complex cropping systems (slash & burn)		90.48	44.35	98.65	814
		Shifting slash & burn activities		98.9	39.39	99.84	112
Forest and semi-natural area	Forest	Moist evergreen forests on the coastal plain land	Low forest on white sand	99.71	43.33	99.9	24
			Coastal forest on rocks	99.8	45	99.9	14
		Flooded forest or swamp		97.37	37.8	100.0	288
	Open space with some/no vegetation	Beach, mud bank, dune		99.78	0	99.99	15
	Degraded natural environment	Forest and shrubs in mutation		93.92	45.53	98.35	602

3.2. Prediction evaluation

Based on a 10-folds cross-validation, we computed the accuracy, sensibility and specificity for the 39 classifiers. All results are described in Tables 1, 2 and 3. Accuracy values vary from 92.44% to 100% with an average value of 98.82%, on the other hand, specificity values vary between 98.02% and 100% and the average is 99.65%. Sensibility varies from 0 to 100% with an average of 70%. Twenty three percent, 36% and 41% of the 39 classes are associated with a sensibility value, respectively, lower than 50% (Table 1), between 50% and 80% (Table 2) and greater or equal to 80% (Table 3).

4. Discussion

The number of induced rules is relatively large. However, the distribution of this number among land cover/use classes is not homogeneous. For instance, we obtained 36 rules for *Forest of the old coastal plain* whilst we had just one rule for the *Forest plantation* class.

From a qualitative point of view, induced rules seem to be consistent with the knowledge on the environment of the study area. Moreover, they are intelligible and interpretable by non-experts. However, some rules appear to be very specific, by covering very few (2 or 3) positive examples whereas the total numbers of positive examples for the classes were large (see rules (4), (5) and (6)).

The predicates *south*, *north* and *west* are not used in the rules whereas the maximum number of literals in the acceptable clauses (set to 6) was not reached. This shows that such predicates were not pertinent for the object discrimination and that we should better characterise the objects by exploiting expert knowledge. In particular, domain ontologies could guide the learning process by specifying the predicates and the learning constraints to use.

We noticed that classes associated with very high sensibility (Table 3) undergoes no or slow changes in time, the knowledge of the land cover type at one time in the past defining, for a large part, the land cover type at present and in the future. It is the case for very anthropised area like *Port area* and *Airport* or for very stable natural land cover types that can not be exploited by

TABLE 2 – Ten-fold cross validation results for land cover/use classes associated with "medium" sensibility values (between 50% and 80%)

Class (level I)	**Class (level II)**	**Class (level III)**	**Precision**	**Sensibility**	**Specificity**	**Total number of positive examples**
Non-agricultural artificial territories	Artificialized green space		99.96	75	99.99	8
	Urbanized areas	Continuous urban base	99.72	67	99.92	42
		Discontinuous urban base	98.88	58.27	99.84	166
		Isolated build	92.44	64.15	98.08	1191
		Multidisciplinary habitat	97.1	67.58	99.56	552
	Industrial zone	Industrial area	99.76	62.5	99.97	40
		Road network	99.44	56.39	99.96	84
	Mines, rubbish construction site	Material extraction	99.05	66.32	99.69	137
Agricultural territories	Arable land	Arable land out of irrigation	99.92	79.55	99.96	12
	Permanent cultivation	Fruit groves	98.35	61.38	99.74	259
	Prairies		98.84	67.78	99.93	243
Forest and semi-natural area	Forest	Mangrove swamp	97.54	67.52	98.67	259
	Degraded natural environment	Degraded forest	97.01	69.78	98.98	483
Free water surfaces	Continental water	River	99.72	62.5	99.89	32

humans due to natural and/or legal constraints, like *Bare rocks, rock savannah*, *Riverine swamp*.

On the other hand, classes associated to low sensibility values (Table 1) seem correspond to continually and rapidly shifting land cover/use types. It is particularly the case for the classes *Construction sites*, *Heterogeneous agricultural areas* corresponding to slash and burn activities, *Beach, mud bank or dune* that is associated with a highly dynamic environment (ANTHONY et al., 2010), *Forest and shrubs in mutation*. For these classes, the information provided by the land cover/use maps seem insufficient in terms of anteriority and/or time resolution.

In our study, we did not exploit the satellite and the aerial photographs information (pixel values, texture, etc.) in order to characterise the objects. Obviously these information would have considerably enrich the background knowledge and improve the results. However, the promising prediction results we obtained show the discriminative power of our approach and that structural knowledge can greatly improved the classification accuracy. Moreover, such knowledge appear to be more robust than these related to image data, which highly depends on sensors and of their associated spatial and spectral resolutions.

Inductive Logic Programming is devoted to symbolic data. Several solutions exist in order to code the numeric data into symbolic one, and the management of numeric information in ILP constitutes a specific research field. The solution we proposed seems to realise a good compromise, during the learning phase, between information loss and generalisation capacity.

Eventually, we adopted the one vs. rest approach for managing more than two land cover/use classes. In further works, we should place ourselves in a multi-class framework in order to manage possible assignment of one object to different classes and to adopt more appropriate validation measures (ABUDAWOOD ; FLACH, 2011).

TABLE 3 – Ten-fold cross validation results for land cover/use classes associated with "high" sensibility values (greater than 80%)

Class (level I)	**Class (level II)**	**Class (level III)**		**Precision**	**Sensibility**	**Specificity**	**Total number of positive examples**
Non-agricultural artificial territories	Industrial zone	Port area		99.99	80	100	5
		Airport		100	100	100	12
Forest and semi-natural area	Forest	Moist evergreen forests on the coastal plain land	Forest on sandy cord	99.69	80	99.83	49
			Forest of the old coastal plain	96.67	80.12	98.02	543
		Wet evergreen forests on laterite	High forest	98.83	82.92	99.27	194
			Low forest	99.97	98.33	99.99	58
		Forest plantation		99.93	86.67	99.97	21
	Shrubby environment	Dry Savannah		99.47	94.49	99.59	164
		Flooded Savannah		99.79	92	99.9	98
	Open space with some/no vegetation	Bare rocks, rock Savannah		100.0	100	100.0	6
	Degraded natural environment	Degraded marshy or flooded forest		99.96	85	100	18
Wetlands	Interior wetlands	Interior marsh		99.51	88.27	99.77	163
		Riverine swamp		99.87	100	99.87	38
	Coastal wetlands	Coastal Marsh		99.97	88.89	99.99	9
Free water surfaces	Continental water	Natural water surface		100	100	100	4
		Fish-pond, artificial pond		99.94	85	99.99	18

5. Conclusion

In this paper we proposed an approach to induce classification rules in order to conceive automatically land cover/use map. Therefore, we implemented a method for automatic extraction of structural knowledge using Inductive Logic Programming. We applied the proposed method to the updating of the land cover/use of the French Guiana littoral.

Results show that the induced rules are intelligible and easy to interpret even by non expert users. In particular, they provide knowledge on structural aspects. A ten-fold cross validation of our classifiers provided very promising results that suggest i) that the accuracy of the automatic classification procedures could be greatly improved by the addition of automatically learned structural knowledge and ii) that we are able to provide a valuable assistance to operators using object-based classification procedures.

Acknowledgements

This work has been performed within the framework of the project CARTAM-SAT (CARtographie du Territoire AMazonien : des Satellites aux AcTeurs - Dynamic mapping of Amazonian Territories : from Satellites to Actors) which is funded by the European Regional Development Funds (FEDER) for French Guiana : agreement number 30492.

Références

ABUDAWOOD, T. ; FLACH, P. A. Learning multi-class theories in ilp. In : *Proceedings of the 20th international conference on Inductive logic programming*. Berlin, Heidelberg : Springer-Verlag, 2011. (ILP'10), p. 6–13. ISBN 978-3-642-21294-9.

ANTHONY, E. J. et al. The amazon-influenced muddy coast of south america : A review of mud-bank-shoreline interactions. *Earth-Science Reviews*, v. 103, n. 3-4, p. 99–121, 2010. ISSN 0012-8252. Disponível em : <http ://www.sciencedirect.com/science/article/pii/S0012825210001200>.

BLOCKEEL, H. et al. Experiments in predicting biodegradability. *Applied Artificial Intelligence*, v. 18(2), p. 157–181, 2004.

CHELGHOUM, N. et al. Fouille de donnees spatiales - approche basee sur la programmation logique inductive. Edition CEPADUES, p. 529–540, 2006.

CORDIER, M.-O. Sacadeau : A decision-aid system to improve stream-water quality. *ERCIM News*, IOS Press, v. 61, p. 37–38, 2005.

DESACHY, J. Connaissances et données exogènes dans l'interprétation d'images satellite : le système expert ICARE. In : *Caractérisation et suivi des milieux terrestres en régions arides et tropicales*. [S.l.] : ORSTOM, 1991, (Colloques et Séminaires). p. 397–427. ISBN 2-7099-1051-9.

DURAND, N. et al. Ontology-based object recognition for remote sensing image interpretation. In : *Proceedings of the 19th IEEE International Conference on Tools with Artificial Intelligence - Volume 01*. Washington, DC, USA : IEEE Computer Society, 2007. (ICTAI '07), p. 472–479. ISBN 0-7695-3015-X.

FONSECA, N. A. ; SILVA, F. ; CAMACHO, R. April - an inductive logic programming system. In : *JELIA*. [S.l. : s.n.], 2006. p. 481–484.

FROMONT, E. ; CORDIER, M.-O. ; QUINIOU, R. Extraction de connaissances provenant de données multisources pour la caractérisation d'arythmies cardiaques. In : BOUSSAID, O. et al. (Ed.). *Fouille de données complexes*. [S.l.] : Cepaduès, 2005, (Revue des Nouvelles Technologies de lĨnformation, RNTI-E-4). p. 25–45.

GOODACRE, J. *Inductive Learning of Chess Rules Using Progol*. [S.l.] : Oxford University, 1996.

GRASS_DEVELOPMENT_TEAM. *Welcome to GRASS GIS*. 1999–2012. `http://grass.fbk.eu/`.

HUDELOT, C. ; ATIF, J. ; BLOCH, I. Fuzzy spatial relation ontology for image interpretation. *Fuzzy Sets and Systems*, v. 159, n. 15, p. 1929–1951, 2008.

LAVRAC, N. ; DZEROSKI, S. *Inductive Logic Programming : Techniques and Applications*. [S.l.] : Ellis Horwood, 1994.

LUU, T.-D. et al. Kd4v : Comprehensible knowledge discovery system for missense variant. *Nucleic Acids Research*, v. 40, n. 1, p. W71–W75, May 2012.

MALERBA, D. et al. Empowering a gis with inductive learning capabilities : the case of ingens. *Computers, Environment and Urban Systems*, v. 27, n. 3, p. 265 – 281, 2003. ISSN 0198-9715.

MUGGLETON, S. Inductive logic programming. *New Generation Computing*, Ohmsha, Ltd., v. 8, p. 295–318, 1991. ISSN 0288-3635.

MUGGLETON, S. Inverse entailment and progol. *New Generation Computing*, v. 13, p. 245–286, 1995. ISSN 0288-3635.

SANTOS, J. et al. Automated identification of protein-ligand interaction features using inductive logic programming : A hexose binding case study. *BMC Bioinformatics*, v. 13, n. 1, p. 162, 2012. ISSN 1471-2105. Disponível em : <http ://www.biomedcentral.com/1471-2105/13/162>.

SRINIVASAN, A. *The Aleph Manual*. March 2007. `http://www.cs.ox.ac.uk/activities/machlearn/Aleph/aleph.html`.

SRINIVASAN, A. et al. Theories for mutagenicity : A study in first-order and feature-based induction. *Artificial Intelligence*, v. 85, p. 277–299, 1996.

VAZ, D. ; COSTA, V. S. ; FERREIRA, M. Fire ! firing inductive rules from economic geography for fire risk detection. In : *ILP*. [S.l. : s.n.], 2010. p. 238–252.

VAZ, D. ; FERREIRA, M. ; LOPES, R. Spatial-yap : a logic-based geographic information system. In : *Proceedings of the 23rd international conference on Logic programming*. Berlin, Heidelberg : Springer-Verlag, 2007. (ICLP'07), p. 195–208. ISBN 3-540-74608-0, 978-3-540-74608-9.