Interpretable algorithms for medical decision support*

João P. Ramos[1], Miguel Tavares[1], Jorge Henriques[1], Jorge Ferreira[2] and Paulo Carvalho[1]

***Abstract*— This work analyses two knowledge-biased strategies of classical interpretable learning methods that take advantage of external knowledge to guide the learning process. This external knowledge takes form in readable and easy to construct IF THEN rules. The case-study is a clinical dataset to predict short-term cardiovascular risk on patients admitted at the hospital with acute myocardial infarction.**

I. INTRODUCTION

This work discusses two different interpretable learning approaches exploiting expert knowledge, in the form of *IF-THEN* rules easily provided by domain experts. The Knowledge-Biased Tree (KB3) algorithm is a white-box, decision tree inducer based on the C5.0 [1]. KB3 uses the external knowledge to select discriminant attributes and cut-off points closer to those existing in the knowledge base. The other method is a Knowledge-biased sUpervised Classifier System (KUCS), a rule-inducer based on UCS [2], a bio-inspired machine learning method. KUCS exploits external knowledge by favouring parents that are in accordance with the existing knowledge and by favouring the specification of attributes known to exist in the knowledge base.

The case study applied in this paper is short-term cardiovascular risk assessment tool. These tools are used to assess the future condition of a patient admitted at a hospital with acute coronary syndrome. The whole dataset was collected by St. Cruz Hospital in Lisbon, Portugal (detailed information in [3]). The strategies under study are compared against the state of the art risk score tool.

II. RESULTS & VALIDATION

All experiments were done using a 10-fold cross-validation methodology. The knowledge-biased strategies are compared against each other and their naive counterparts. Overall, there is not much of a difference between strategies. In the 'UCS vs KUCS' (Table I) experiment we observe that all the slight variations have no statistical significance. This gives us reasons to believe the information from the external knowledge 'disagrees' with what results from mining the data and any influence the external knowledge might produce is lost and forgotten through the various epochs.

In contrast, the 'C5.0 vs KB3' presents a quirk in the behaviour. The algorithms reverse their *sensitivity* and *specificity* values, meaning that the KB3 is more sensible

TABLE I

ALGORITHMS COMPARISON.

	UCS			KUCS		
	G_MEAN	SE	SP	G_MEAN	SE	SP
infarctionA	0.41	0.27	0.71	0.42	0.27	0.71
infarctionB	0.41	0.26	0.77	0.42	0.29	0.78
	C5.0			KB3		
	G_MEAN	SE	SP	G_MEAN	SE	SP
infarctionA	0.46	0.32	0.68	0.49	0.65	0.40
infarctionB	0.51	0.38	0.75	0.60	0.65	0.57

	GRACE		
	G_MEAN	SE	SP
infarctionA	0.63	0.52	0.78
infarctionB	0.61	0.49	0.78

to positive cases, but at the same time is producing more false alarms. Since all rules have a 'positive' outcome, the external knowledge may be over-steering the learning process towards the positive class.

Between the two biased strategies the results are disparate. As previously observed, the KB3 has an undesirable high false positive rating. On the other hand, KUCS is more conservative on triggering false positives but it is at the expense of its sensitivity. Overall, KB3 performs better in both cases - a higher *G_MEAN* (p-value < 0.05).

Despite the weak results, in comparison with the currently widely used risk score - GRACE - the results are not so different. The GRACE score has a low *sensitivity* for the specific threshold used and a possible adjustment might increase the false alarms. Moreover, any of the other solutions should be sought-after as they are more flexible and can easily learn from new unseen data as they have either means for online learning or can be easily constructed.

III. CONCLUSIONS

In both strategies we can conclude that external knowledge is yet to be used in an efficient manner. Specifically, KUCS has a similar performance to its naive counterpart, UCS. On the other hand, KB3 seems to be influenced by the used external knowledge, increasing the false alarms rate. In the clinical domain, false alarms can induce alert fatigueness, diminishing the system's trustworthiness.

REFERENCES

[1] R. Quinlan, "Data mining tools see5 and c5. 0," http://www.rulequest.com/see5-info.html, 2004.

[2] E. Bernadó-Mansilla and J. M. Garrell-Guiu, "Accuracy-based learning classifier systems: Models, analysis, and applications to classification tasks," *Evolutionary Computation*, vol. 11, pp. 209–238, 2003.

[3] S. Paredes, "Integration of different risk assessment tools to improve the event risk assessment in cardiovascular disease patients," Ph.D. dissertation, University of Coimbra, 2012.

*This work was supported in part by EU Project LINK (H2020-LINK 692023)

[1]CISUC, Departament of Informatics Engineering, Coimbra University, Portugal {jpramos, mtavares, jh, carvalho}@dei.uc.pt

[2]Cardiology Department, Santa Cruz Hospital, Lisbon, Portugal jorge_ferreira@netcabo.pt