JYVÄSKYLÄN YLIOPISTO
UNIVERSITY OF JYVÄSKYLÄ

This is a self-archived version of an original article. This version may differ from the original in pagination and typographic details.

Author(s): Subirats, L.; Fort, S.; Hernández, C., Pérez; L., Vesisenaho, M.; Nousiainen, T.; Peltonen, M.; Miakush, I.; Sacha, G.M.

Title: Intelligent Tutoring System in Archaeology

Year: 2019

Version: Published version

Copyright: © 2019 Association for the Advancement of Computing in Education (AACE)

Rights: In Copyright

Rights url: http://rightsstatements.org/page/InC/1.0/?language=en

Please cite the original version:

Subirats, L., Fort, S., Hernández, C., L., V., Nousiainen, T., Peltonen, M., Miakush, I., & Sacha, G.M. (2019). Intelligent Tutoring System in Archaeology. In T. Bastiaens (Ed.), Proceedings of EdMedia + Innovate Learning 2019. Association for the Advancement of Computing in Education (AACE). https://www.learntechlib.org/primary/p/210074/

Intelligent Tutoring System in Archaeology

Laia Subirats, Santiago Fort
Eurecat – Centre Tecnològic de Catalunya
Spain
laia.subirats@eurecat.org, santi.fort@eurecat.org

Cristo Hernández
Departamento de Didácticas Específicas, Universidad de La Laguna
Spain
chergomw@gmail.com

Leopoldo Pérez
IPHES – Institut Català de Paleoecologia Humana i Evolució Social and Universitat Rovira i Virgili
Spain
ljperez@iphes.cat

Mikko Vesisenaho, Tuula Nousiainen, Marika Peltonen, Iryna Miakush
Department of Teacher Education
University of Jyväskylä
Finland
mikko.vesisenaho@jyu.fi, tuula.j.nousiainen@jyu.fi, marika.p.peltonen@jyu.fi, iryna.miakush@gmail.com

G. M. Sacha
Universidad Autónoma de Madrid
Spain
sacha.gomez@uam.es

Abstract: A method that uses artificial intelligence for the taxonomical characterization of bone remains in archaeological sites is shown. The main goal of this method is to help students and archaeologists in the classification of samples in order to improve the efficiency of their tasks during the campaigns at archaeological sites. The development of the system implies several steps: training of classification algorithms, development of a user-friendly interface and implementation of gamification techniques to improve learning motivation and the efficiency of the system in the learning process. This system will help archaeology students to classify new samples. In this prototype, by introducing characteristics of a sample, the system answers with a possible animal "Family" related to the sample. This answer can be used by archaeologists as a first clue to determine the species recovered, or by students as additional information to help them in their learning process and self-regulation learning phase.

Introduction

The term *artificial intelligence (AI)* was coined by John McCarthy in 1956. Since then, many initiatives for applying AI in the archaeology domain have been developed. Puyol (1999) has presented some areas that could be applied to archaeology: Knowledge Discovery in Databases (KDD), Visual Information Management (VIM) and Multi-Agent Systems (MAS). Barcelo (2009) summarizes some of these initiatives, such as visual and non-visual analysis in archaeology (Barcelo, 2010), shape analysis, texture and compositional analysis, spatiotemporal analysis and an automated approach to historical and social explanation. In addition, Barcelo and Bogdanovic (2015) explain some examples of mathematics of ancient artifacts, mathematics of archaeological time and space, and modelling social action in the past. Azkarate (2018) also introduces the use of clustering analysis as a first step in quantitative techniques in archaeology.

In the past, *intelligent tutoring systems (ITS)* and *interactive learning environments (ILEs)* were seen as the two options of applying AI to education. Nowadays, the possibilities are seen as a continuum with a controlling tutor at one end, and a completely controlling student at the other (McArthur, 2005).

The goal of this work is to develop a system for automatic evaluation of bone fragments that can be used for both archaeologists and students. In archaeological sites, sometimes the circumstances of the campaign imply that the experts of a specific field are not available or even are not participating in the campaign at the moment when their expertise is needed. For example, in a long campaign, the zooarchaeologists can be participants in the first half of the campaign and finish their job at their workplaces during the second half. In this case, participants of the second half of the campaign cannot benefit from their knowledge. In these cases, having a tool that can identify bone fragments in real time, or at least give a first clue, could be of great help.

In the case of students, it can be helpful in two different ways. First, the automatic characterization of bone fragments can be used for students when working on a specific fragment that must be identified. In this case, the system can be a helpful substitution of the zooarchaeologist, who may not have time to answer all the requests made by students in the case of large groups. An automatic tutoring system can be very helpful in this case to improve the efficiency of the work in the archaeological site. This is a very important fact to take into account in archaeological campaigns since sometimes the time for working in the field is strictly limited. A second potential application for our system is related to learning or self-evaluation since students can use it to evaluate their accuracy when working on bone fragments. In this case, the system can be used as a check tool or even as a final evaluation platform.

Thus, this article shows a method that uses artificial intelligence to develop an efficient platform for the taxonomical characterization of bone fragments in archaeological sites. The main goal of this method is to help students and archaeologists in the classification of samples in order to improve the efficiency of their tasks during the campaigns at archaeological sites and/or their subsequent laboratory analysis. The development of the system implies several steps, which involve 1) initially the training of classification algorithms (Random Forests, Support Vector Machines, Neural Networks, Naive Bayes and K-nearest neighbors (Cady, 2017)), 2) the development of a user-friendly interface and 3) the implementation of gamification techniques to improve learning motivation and thereby the efficiency of the system in the learning process.

Development of the Classification Algorithms

For the testing of our method, we have used data from the archaeological site of El Salt in Alcoy (Alicante, Spain) (Perez et al., 2015, 2017) in the field of archaeozoology. The dataset we used had 3406 instances of archaeological samples and 64 attributes (the last one is the predicted class). In Table 1, a descriptive analysis of the sample is shown for the most relevant parameters, according to the classification algorithm with best performance in the coarse, medium and fine granularities. Numerical attributes are described with count (number of not null values), mean, standard deviation (std), minimum (min), maximum (max), quartile 1 (Q1), quartile 2 (Q2 or the median) and quartile 3 (Q3). Categorical attributes are described with count (number of not null values), unique values, top values and frequency of the top value (freq). To avoid categories with too small representation in the sample, we have created a class called “unknown” where all these classes are included. The final categories are 420 bovidae, 516 cervidae, 240 equidae, 2164 leporidae and 66 unknown.

Attribute	**Count**	**[Mean, Std]**	**[Unique, Top, Freq]**	**[Min, Max]**	**[Q1, Q2, Q3]**
Width	2036	[14.82, 8.97]	-	[0.99, 71.01]	[8.45, 13.43, 19.32]
Bone	3403	-	[93, T, 374]	-	-
Thickness	714	[4.88, 2.87]	-	[0.46, 16.52]	[3.07, 4.59, 6.4]
Length	2492	[36.25, 28.28]	-	[2.42, 284.96]	[15.94, 27.20, 47.80]
Bone fragment	3242	[189.63, 204.90]	-	[1, 555]	[50, 111, 500]
Anatomical group	3399	-	[11, Mp, 1275]	-	-
Long bone circumf.	1378	[2.52, 1.10]	-	[1, 5]	[2, 2, 4]
X	1187	[1864.52, 62922.28]	-	[0, 2167891]	[8.51, 29, 67.75]
Y	1186	[1272.85, 30223]	-	[0, 803982]	[11.80, 24.50, 63]

Z	1232	[403.04, 9767.16]	-	[-143.8, 342938]	[-0.156130, 166.2, 192.6]
Manganeso	1416	[1.59, 0.83]	-	[1, 5]	[1, 1, 2]
Concretion	1177	[1.80, 1.03]	-	[1, 5]	[1, 1, 2]

Table 1: Exploratory Data Analysis of the dataset

There were 67% instances for the training set and 33% instances for the validating set. A pipeline with standard scaler with a Synthetic Minority Over-sampling Technique (SMOTE) or Adaptive Synthetic (ADASYN) method was used, together with 10-fold cross-validation.

The parameter that is wanted to be maximized is f1-score_macro (it calculates metrics for each label, and finds their unweighted mean; this does not take label imbalance into account). The Random Forest has been used with parameters n_estimators [100,500,700], max_features [auto, sqrt, log2], class_weight [balanced,None] which are the number of trees of the forest, the number of features to consider when looking for the best split and weights associated with classes respectively. Regarding the Support Vector Machine (SVM), the parameters have been kernel [linear,rbf,poly], gamma [1e-2,1e-3,1e-4] and C [1, 10, 100], which are the kernel to be used in the algorithm, the kernel coefficient, and the penalty parameter C of the error term. Neural Networks with parameters solver [lbfgs, sgd, adam] and alpha [1e-4,1e-5] which are the solver for weight optimization and L2 penalty (regularization term) parameter. And finally, the parameters used in the k-nearest neighbors (KNN) where n_neighbors [3,5,7,9] which are the number of neighbors used. Naive Bayes was used without parameters. These algorithms have been implemented using the Python library scikit-learn and the tutoring system has been implemented using Flask.

The results of the experiments demonstrated that random forest is the method that gives the best performance. The generated confusion matrix of this method is shown in Figure 1.

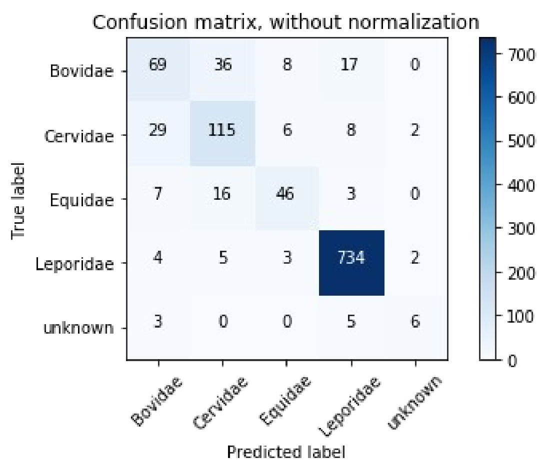

Figure 1: Confusion matrix of the random forest without normalization.

We have found that the first 10 more relevant attributes for the classification are width, bone, thickness, length, bone fragment, anatomical group, long bone circumference, X, Y and Z. The best configuration for the random forest was: {'classifier__max_features': 'auto', 'classifier__class_weight': None, 'classifier__n_estimators': 500}

Development of the Intelligent Tutoring System's Interface

The interface of the intelligent tutoring system prototype has been developed in a user-friendly form in order to be used by people without advanced knowledge in numerical algorithms. As mentioned previously, the target of this platform is students or archaeologists who will get new data of bones or fragments in archaeological sites and need confirmation of the species most probably related to the fragment. In the case of students, this tool will be helpful for them to automatically get a possible answer without having any help from the archaeologists. In the case of archaeologists, the tool can be helpful in the cases where it is not easy to find the species as the tool will give them a possible answer that can be considered a first step or a clue to start the analysis of the fragment. In the interface, shown partially in Figure 2, we can see open text boxes that allow users to include all the relevant data from the fragment.

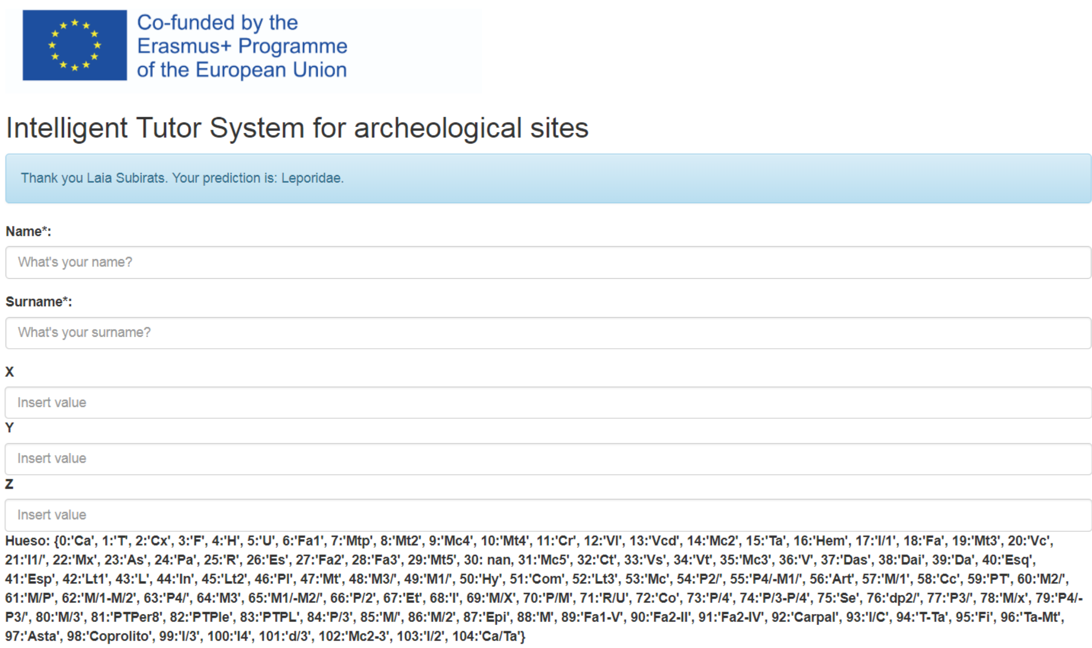

Figure 2: Form for archaeology students to introduce the data and receive the classification of the sample in the coarse granularity.

When all possible data has been included, the algorithms included in the application will give a possible answer about the "Family" that is related to the fragment. In the application, we have also included the option of identifying the user in order to evaluate or identify their results. This identification is important both for archaeologists and for students. In the case of archaeologists, it is important because the results obtained for every fragment can be used in the future in other scientific publications and the authority must be clear. In the case of students, the identification allows the students to follow their trajectory and learning improvement. On the other hand, it is also necessary for the evaluation process.

Next Step: Development of Gamification Techniques

When using the platform as a self-evaluation or learning tool, we must design a motivating system that increases the correct and frequent use of the platform by the students. An earlier pilot study (Comas-Lopez, 2018) conducted with a similar tutoring system in a different higher education context (chemistry and chemical engineering) identified a set of main challenges pertaining to student motivation. The first one is the need to increase the students' attention to task, which means that instead of only cursorily going through the tasks when practicing with the tutoring system, they should focus on the items more closely. Another challenge was related to enhancing the regular use of the tutoring system already from the beginning of the course instead of only practicing with it a few days before the exam.

To this end, the use of *gamification techniques* is extremely relevant. Gamification refers to using game thinking and game elements for turning a non-game activity into a game, making it more attractive and motivating (Deterding et al., 2011; Farber, 2015; Kapp, 2012). Some typical examples include elements such as rewards, points, badges, and leaderboards. However, specific gamified applications and elements may have a different – even opposite – impact on different individuals (e.g., Hamari & Tuunanen, 2014; Lopez & Tucker, 2019), which makes it difficult to design one solution that fits all.

Therefore, we propose an approach where we first implement a "light version" of gamification elements to be used by all students, in addition to which we can provide a set of optional tools tailored to different types of users (e.g. those motivated by achievement, those motivated by social interaction, and those motivated by exploring the environment; see e.g. Hamari & Tuunanen, 2014). In this case, we suggest that the light version be principally based on badges, as they allow us to tap into many of the problematic elements identified in the pilot study. To mention some examples, we can provide badges rewarding the regularity of use (e.g., on a daily or weekly basis), prediction accuracy (i.e., performance), and the number of times all the tasks are completed (encouraging frequent rehearsal). A pilot study can be run using the light version, and in case some gaps are identified, additional gamification elements can be targeted specifically to fill them. For example, optional leaderboards may be made available to those motivated by competition, and some advanced visualization of progress to those who value exploration.

Conclusions

In this article, we have proposed a tutoring system that can be a useful tool to archaeologists or students in order to help them to classify bone fragments. The classification system has been developed from a sample of 3000 correctly identified bone fragments from "El Salt" archaeological site in Alcoy, Spain. For the classification system, we have tested five different supervised classification algorithms and three different classification distributions (coarse, medium and fine granularities). We have found that better results have been obtained by the random forest method, with an accuracy and f1-score of both 0.86 in coarse granularity (with samples including bovidae, cervidae, equidae, leporidae and unknown). Results from the classification algorithms indicate that the most relevant attributes for the classification are width, bone, thickness, length, bone fragment, anatomical group, long bone circumference, X, Y and Z. In order to make the system useful, we have developed a user-friendly interface that is connected to a random forest classifier. This interface is composed of a questionnaire to introduce all the possible data from the bone fragment, including 63 possible attributes. In order to improve the efficiency of the method, we have suggested an approach where both generic and personalized gamification techniques (such as badges, leaderboards, and progress visualization tools) are applied in order to encourage frequent and regular use of the tutoring system. This approach is motivated by previous results showing that there was a need for solutions that encourage students to focus on the tasks and to take up studying earlier than at the very end of the course. In our future work, the tutoring system will be tested in the archaeological site of "El Salt" and the usefulness of the granularities will be compared.

References

Azkarate, A., García-Gómez, I., & Mesanza-Moraza, A. (2018). Cluster Analysis: A First Step in Quantitative Techniques in Archaeology of Architecture. *Arqueología de la Arquitectura*, 15, e066.

Barcelo, J. A. (2009). *Computational Intelligence in Archaeology.* Hershey, PA: Information Science Reference.

Barceló, J.A. (2010). Visual Analysis in Archaeology: An Artificial Intelligence Approach. In A. M. T. Elewa (Ed.), *Morphometrics for Nonmorphometricians* (pp. 93–156). Berlin: Springer.

Barcelo, J. A. & Bogdanovic, I. (Eds.) (2015). *Mathematics in Archaeology.* Boca Raton, FL: CRC Press.

Cady, F. (2017). *The Data Science Handbook.* Hoboken, NJ: Wiley.

Comas-Lopez, M., Hincz, K. P., Gamez, A., Yañez-Mo, M., & Sacha, G. M. (2018). Adaptive Tests as a Supporting Tool for Self-evaluation in Theoretical and Practical Contents in Biochemistry. In F. J. García-Peñalvo (Ed.), *Proceedings of the Sixth International Conference on Technological Ecosystems for Enhancing Multiculturality* (pp. 180–184). New York, NY: ACM.

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From Game Design Elements to Gamefulness: Defining "Gamification". In *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments* (pp. 9–15). New York, NY: ACM.

Drennan, R. D. (2009). *Statistics for Archaeologists: A Common Sense Approach* (2nd ed.). New York: Springer.

Farber, M. (2015). *Gamify Your Classroom. A Field Guide to Game-based Learning.* New York: Peter Lang Publishing.

Hamari, J. & Tuunanen, J. (2014). Player Types: A Meta-synthesis. *Transactions of the Digital Games Research Association, 1*(2), 29–53.

Kapp, K. (2012). *The Gamification of Learning and Instruction: Game-Based Methods and Strategies for Training and Education.* Hoboken, NJ: Wiley.

Lopez, C. E. & Tucker, C. S. (2019). The Effects of Player Type on Performance: A Gamification Case Study. *Computers in Human Behavior, 91,* 333–345.

McArthur, D., Lewis, M., & Bishary, M. (2005). The Roles of Artificial Intelligence in Education: Current Progress and Future Prospects. *Journal of Educational Technology 1*(4), 42–80.

Pérez, L. J., Machado, J., Hernández, C. M., Morales, J. V., Brugal, J. P., & Galván, B. (2015). Arqueozoología y arqueoestratigrafía del yacimiento de Salt (Alcoi, Alicante): Contribución metodológica para el análisis del registro faunístico contenido en palimpsestos arqueológicos del paleolítico medio. In A. Sanchis & J. L. Pascual (Eds.), *Preses petites i grups humans en el passat. II Jornades d'arqueozoologia* (pp. 223–244). València: Museu de Prehistòria de València.

Pérez, L. J., Sanchis, A., Hernández, C. M., & Galván, B. (2017). Paleoecología de macromamíferos aplicada a los conjuntos zooarqueológicos de El Salt y el Abric del Pastor (Alcoy, Alicante). In *Interaccions entre felins i humans. III Jornades d'arqueozoologia* (pp. 327–354). València: Museu de Prehistòria de València.

Puyol-Gruart, J. (1999). Computer science, artificial intelligence and archaeology. In J. A. Barceló, I. Briz, & A. Vila (Eds.), *New Techniques for Old Times. CAA98. Computer Applications and Quantitative Methods in Archaeology. Proceedings of the 26th Conference, Barcelona, March 1998 (BAR International Series 757)* (pp. 19–28). Oxford: Archaeopress.

Shennan, S. (1988). *Quantifying Archaeology.* Edinburgh: Edinburgh University Press.

Wenger, E. (1987). *Artificial Intelligence and Tutoring Systems: Computational and Cognitive Approaches to the Communication of Knowledge.* San Francisco, CA: Morgan Kaufmann.

Acknowledgements

The work reported in this paper has been carried out as part of the ADeAPTIVE (Advanced Design of e-Learning Applications Personalizing Teaching to Improve Virtual Education) project with the support of the Erasmus+ programme of the European Union (grant number 2017-1-ES01-KA203-038266). This work was also part of the project Explora "Application of Soft Computing techniques and 3D modeling in massive data processing in Archeology" that was supported by the Spanish Ministry of Science and Innovation (Grant TIN2015-72682-EXP).