Discovering the dinosaur

Luis Castillo
Dept. Computer Science and AI
University of Granada
Spain
L.Castillo@decsai.ugr.es

Juan Fdez-Olivares
Dept. Computer Science and AI
University of Granada
Spain
Faro@decsai.ugr.es

Francisca Sánchez-Carrascosa
Acceptare Psicología
Spain
fsc.psicologa@gmail.com

***Abstract*—Talking about leadership or collaboration patterns inside a team of people is a tough task, they are hard to define and they are slippery to detect with questionnaires or peer ratings. However, these team dynamics leave a footprint in their daily activity record which could be used to infer interesting properties and try to determine these team dynamics, and even leadership patterns, from a different, non intrusive, point of view. It is like figuring out how a dinosaur was just by analyzing its footprints found in the mud. The result is really surprising.**

Keywords – Teamwork, Leadership, Learning analytics

I. About paleontology

Paleontology deals with the scientific reconstruction of living beings and ecosystems of a distant past from their known evidences found in fossils nowadays, like footprints or remains. Since only partial evidence is usually found, the rest of these animals and ecosystems are deduced by, let say, scientific speculation. In the same way, this work tries to reconstruct the behavior of a team of developers (students during their practice laboratory) from the footprints found in their records of project meetings, under the Scrum methodology.

This reconstruction aims at finding the type of team work, collaboration patterns or leadership among the members of the team, in order to identify good behaviors or potential dysfunctions. Among the features worth to be promoted we may find a high participation of all members, a high collaboration in all tasks, and a good balance of the effort among all members. On the contrary, some of the dysfunctions recurrently cited in the literature [10], [12], [13] are dominance of one team member (very often the leader/coordinator) who imposes his way of doing things to the rest of the team, lack of collaboration or high individualism, resistant members (couch potatoes) who participate very little or none. The role of the leader or coordinator of the team [12] might also determine the dynamics of the team [5]. However, in the literature there is no clear consensus [6] on how to measure leadership capacity and the types of leaders that exist and their influence on the style of teamwork. The theories developed range from the analysis of the leader in the context of the team, the relationships between members and the tasks to be developed to the analysis of the leader's behavior [10], the latter being one of the most widely accepted models. It states the following three types of leaders. Authoritarian leader: the team leader makes the decisions and the team executes them. Democratic leader: any work plan is discussed within the team, something that the leader encourages. Laissez-faire leader: the leader does not exercise any authority and gives total freedom of action to the team members, with little or no intervention. Usually, all this information might be obtained by presenting surveys or peer ratings to the teams of students, but this is very intrusive and tedious for them, with rather long and subjective questionnaires.

In this work, a non-intrusive method is suggested based on very simple rules just by digging into the historic records of Scrum activity. Final results are 100% validated by well known unsupervised and supervised artificial intelligence classification methods. This is the work plan:

1) During the last 7 years, let the students (118 teams and 500+ students) do their work, while recording their historic data. This would be like the footprints in the mud. Extract a appropriate set of key indicators and build a dataset in R.
2) Apply R packages **NbClust()** and **ClValid()** [1] to get to know which is the optimal number k of categories among these 118 teams (Table I).
3) Apply R Package **kmeans()** [7] to divide the 118 teams into k categories (Table II).
4) Use this classification to feed R package **C5.0()** [9] to describe these k categories and get to know their strengths and their dysfunctions (Figure 6).

II. Description of the fossils

In Computer Science Grades, regarding the development of traversal competencies like team work or leadership, grouping the students into teams during practice laboratories is very usual. Even more, the role of leader, or coordinator, could also be designated, among other roles [12] to interface with the teacher, and conduct the work of the team in some way. In many cases, the use of supporting methodologies like Scrum [8] may benefit the self organization and evolution of these teams of students [3] and they allow to keep a record of activity which might be imprinted with part of the dynamics within the team.

A. *Scrum in a nutshell*

The practice laboratory of the course Agent Based Development in the 4th year of the Computer Science Grade in the University of Granada is organized into teams of students following a Project-Based Learning approach and based on the use of Scrum to organize the teams.

We have a set of teams $T = \{t_1, t_2, \ldots, t_n\}$ where every team t_i is made up of a set of people $t_i = \{p_1, p_n, \ldots, p_m\}$ with m varying from 3 to 6. There is a leader $p_j \in t_i$ who is in charge of conducting the team. The laboratory is divided into a set of days named **sprints** $S = \{d_1, d_2, \ldots, d_s\}$ with s ranging from 15 to 21. During a sprint S, each team t_i has to carry out a software development project with a set of features named stories $B = \{s_1, s_2, \ldots s_k\}$ with k varying depending on $|t_i|$ and $|S|$. This set B is commonly known as the **sprint backlog**. Well, all the teams meet every week with the teacher and build an activity record R with information about the assignment of every p_i to any s_j during a certain time t_r in the day d_l

$$R = \{r_1, r_2, \ldots r_\infty\} \text{ where } r_k =< d_l, p_i, s_j, t_r >$$

This is a record of activity [2] kept during all this time, with different generations of students, different PBL projects, and different durations.

B. *The meaning of fossils: key indicators*

From the record R a set of higher level indicators, might be defined in order to measure certain collaboration habits, both good and bad, at individual level or at team level. Although they are formally defined in [4], an informal description is also given here. In order to be able to compare records from one year to another, with different team sizes, durations and features, all these indicators have been normalized in the interval $[0, 1]$ and can be applied to any member p of the team X_p. Two singular values will be highlighted, one applied to the leader X_L, and the other applied to the median of the team X_T.

Heading H. It is the percentage of the sprint in which a team member has headed the effort, that is, the amount of time dedicated to the project. A heading of $H_L = 0.75$ for the team leader means that s/he has dedicated a daily time to the project higher than his team companions during 75% of the sprint S. On the contrary, a heading $H_T \approx 0.5 \approx H_L$ would mean that, the effort seems balanced among the team.

Overload O. It is the intensity of H, that is, when a team member p has a non empty heading $H_p > 0$, that is, s/he has lead the effort of the sprint during a certain period of time, the overload O_p represents the difference between his/her amount of time and the average time of the team. An overload $O_p = 0.10$ means that, during his/her heading, s/he spent 10% more of time than the average of the team.

Participation P. It is the percentage of stories in which a person p_i has participated, no matter the amount of time devoted to it.

Collaboration C. It is the percentage of team members that take part of the same story, averaged along all stories.

Dominance D. It is the percentage of stories in which a person p_i has participated in a story and, in addition to this, he has been the person who spent more time of all the team in the same story.

Individualism I. It is an extreme of dominance, that is, it is the percentage of the stories that a person p_i has developed exclusively by himself, without the intervention of any other team member.

Table IV (up, in black color) shows a summary of these indicators. Why these indicators and no other ones? Simply because they can detect and track the balance of the effort, cooperation patterns and dominance or individualism.

III. Analysis of fossils

A. *Descriptive analysis of team dynamics*

Without doing much further analysis, in all this time and along all these teams, the plot shown in Figure 1 shows something important since it assigns the attitude of the leader an apparently relevant influence. In this case, the commitment of the leader to head the effort of the team ($H_L \gtrapprox 0.5$) or not, seems to polarize the behavior of the teams into two clear segments with barely nothing in between, that is, almost 90% of cases are under 0.4 or above 0.6.

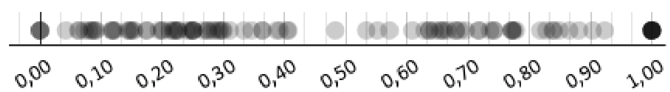

Fig. 1. Frequencies of values of H_L. The darker the dots, the more frequent are these values

Digging deeper, Figure 2 (left) shows a clear correlation: the more participant a leader is (P_L in the X-axis), the more collaborative (C in the Y-axis) and less individualist (I_T as the radius) is the whole team. Although there cannot be proven that there is a cause-effect relationship, it is clear that both indicators are strongly related. In particular the Pearson correlation between P_L and C shows a value $r = 0.736$ with $p = 2.2E - 16$ at a 95% of confidence.

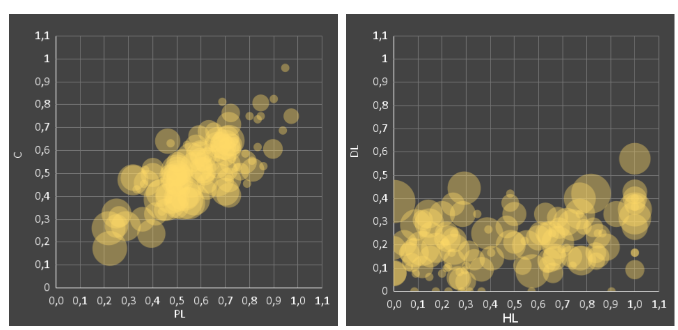

Fig. 2. (Left) Correlation between P_L in the X-axis and C in the Y-Axis, with I_T as dthe radius of the circles. (Right) Correlation between H_L in the X-axis and D_L in the Y-Axis, using I_T as the radius of the area.

Although a high heading of the leader might suggest a dominance over the whole team, it does not have to be this way. Figure 2 (right) shows another correlation between the heading of the leader H_L and the dominance of the leader D_L enhanced with the individualism of the team I_T as the radius of the area and, it does not clearly show any correlation between them.

Since the record R also keeps the sizes of the teams, the Figure 3 shows something that was also described, almost exactly, in [12] but also quantified in [3]: the larger the teams, the less collaborative they are. The most collaborative teams are those of 4 individuals, which rapidly degrades as the size increases. In this case, Pearson's correlation is moderately strong but meaningful $r = 0.436$ with $p = 8.13e - 07$ and a 95% of confidence.

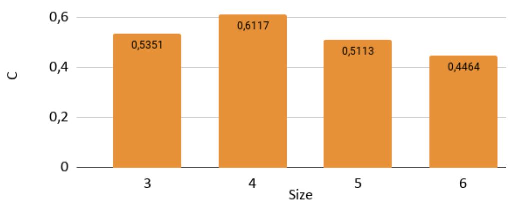

Fig. 3. Average collaboration C found along the different team sizes

Another quite obvious finding is shown in Figure 4 (up). team collaboration C and team individualism I_T are antagonists to each other, but the real finding is that this relation depends on the size of the team, showing more intensity as the size decreases (please note the logarithmic scale in the Y-axis and the coloring of the teams). One could say that smaller teams are more sensible to individualism than larger teams.

Continuing with individualism patterns, Figure 4 (down) shows a clear correlation between an attitude of dominance of the leader D_L and the individualism of the team I_T which does not seem to depend on the size of the team ($r = 0.413, p = 3.26e - 06$).

IV. Quantitative analysis of leadership patterns

In the introduction, we mentioned that there seem to be three types of leaders/teams [5], [10]. Now, let us see if this classification also makes sense in the whole historic data. Henceforth, the 11 indicators gathered would represent each of the 118 teams in a normalized, $[0, 1]$, 11-dimensional space and we will try to prove that a plausible categorization can be found. In order to reduce human bias in this clustering process, a non-supervised clustering method is used named Kmeans [7], [11] which is available in R with the function **R::kmeans()**.

The clustering technique Kmeans gets the raw dataset and tries to find the best k clusters where k is a free parameter given beforehand. Therefore, it is also relevant knowing what the best choice for k is and then finding the best k clusters. This is done with the R packages **R::clValid** and **R::NbClust**. Both packages analyze multiple indexes, each of which propose a certain k as the most plausible choice and the result is shown in Table I. Besides, Table II shows the assignment to each of the 3 categories.

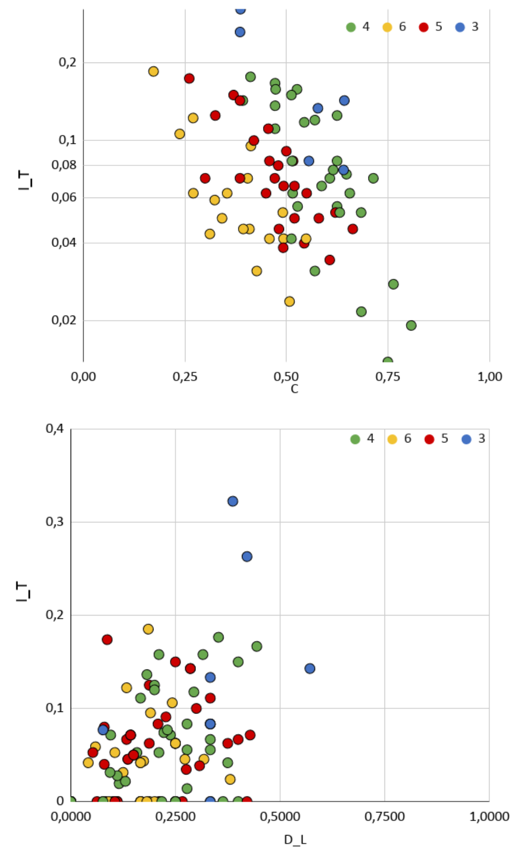

Fig. 4. (Up)Correlation between the collaboration C in the X-axis and individualism I_T in the Y-axis of the same team, segmented by team sizes (coloured). (Down) Influence of a dominant leader D_L in the X-axis into the individualism of the team I_T in the Y-axis segemted by team sizes (coloured)

$k =$	3	4	5	6	7	8	9	10
points	12	5	1	0	2	3	2	5

TABLE I
BEST CHOICE OF THE OPTIMAL NUMBER OF CLUSTERS k, UPON BOTH R PACKAGES **R::clValid** AND **R::NbClust**

A. *Leaving out borderline teams*

Then, while maintaining the $k = 3$ clustering, the set of 118 teams is going to be filtered just to leave out weakly connected teams. Therefore, having $k = 3$ clusters or categories, we are going to keep only those strong components with maximal connectivity, that is, those maximal subsets that always appear connected between them in all clusters found from $k = 1$ to $k = 3$. They might be considered as "reliable" components, and all teams that do not meet this condition are left out of the problem. Interestingly, there are only 3 of these weakly connected teams: $\{17, 35, 115\}$. Table III shows that, in all calculated indices, the new partition with $n = 115$ fits much better than the one with $n = 118$.

B. *Shaping the dinosaur*

Now we start to see the dinosaur who imprinted the historical data. Figure 5 shows the 115 selected teams within the

TABLE II
CATEGORIES FOUND WITH KMEANS FOR $k = 3$

Unfiltered clusters, k=3, n=118

Cluster1	Cluster2	Cluster3
3 13 14 15 17 19 21 22 26 28 30 46 50 52 55 60 70 77 78 79 82 87 89 91 97 98 100 103 105 106 110 113 114	5 6 7 10 11 16 20 24 25 31 33 34 36 38 39 40 41 42 43 44 47 48 51 53 54 56 57 58 59 61 62 63 64 65 66 67 68 71 73 74 75 83 84 85 86 88 90 93 96 99 101 104 107 108 109 112 115 118	1 2 4 8 9 12 18 23 27 29 32 35 37 45 49 69 72 76 80 81 92 94 95 102 111 116 117

TABLE III
QUALITY METRICS APPLIED WITH THE WHOLE SET ($n = 118$) AND AFTER REMOVING WEAKLY CONNECTED TEAMS ($n = 115$). THE UPPER PART OF THE TABLE SHOW INDICES WHICH ARE BETTER AS THEY INCREASE. THE LOWER PART ARE INDICES WHICH ARE BETTER AS THEY DECREASE

		$k = 3$	
	better	$n = 118$	$n = 115$
$\frac{BSS}{TSS}$	$\nearrow$	43.9	44.4
Dunn	$\nearrow$	0.105	0.113
Silhou.	$\nearrow$	0.311	0.324
APN	$\searrow$	0.100	0.086
Connect.	$\searrow$	28.289	24.504
AD	$\searrow$	0.475	0.465
ADM	$\searrow$	0.069	0.059
FOM	$\searrow$	0.116	0.115

plane formed between H_L in the X-axis and H_T in the Y-axis. It may be seen how clear these clusters are with little overlap between them.

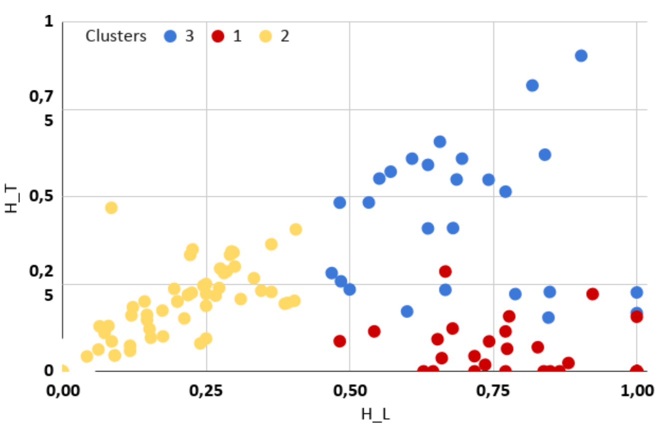

Fig. 5. $k = 3$ clusters found in the cutting plane between H_L in the X-axis and H_T in the Y-axis-

However, before start interpreting these results, we are going to use these $k = 3$ clusters as input to the well known supervised classification procedure **R::C5.0()** whose result is shown in Figure 6 in the form of a decision tree with an error of 0% within just 6 iterations. That is, if $H_L < 0.406$ then the cluster is 2 (yellow in Figure 5). Otherwise, if $H_T <= 0.162$ then the cluster is 1 (red). Otherwise, if $H_T > 0.343$ then the cluster is 3 (blue) and so on. In addition to this, Table IV shows some statistics of these clusters which will help understand all these results.

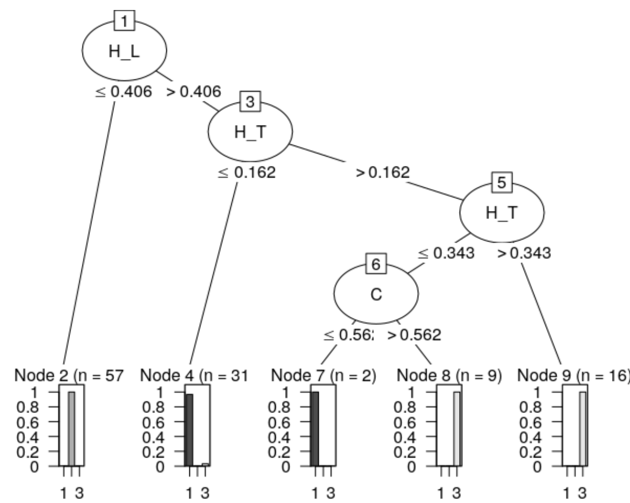

Fig. 6. Decision tree obtained in R with the package **C50**.

C. About Cluster 2 (yellow)

This cluster is amazingly well described by a pretty simple absolute rule refered to the leader, $H_L < 0.406$, whose fitness is 100% and it also matches the left hand side polarization shown in Figure 1. However, we could complement this rule with some additional information just by looking at the statistical data in Table IV and the distribution of plots in Figure 5.

1) The heading of all yellow leaders is the lowest one, completely underneath the blue and red leaders together

$$max(H_L^y) < min(H_L^{r,b})$$

2) More than 80% of yellow leaders head their sprint during less than $1/3$ or even nothing at all.

$$0 <= H_L^y < 0,33$$

3) Although it is not numerically documented here (due to lack of space), in 100% of yellow teams, there is at least one member of the team with a heading greater than the own leader

$$\forall t \in Cluster^y, L \in t, \exists p \in t, H_L^y < H_p^y$$

4) Yellow teams (other than the leader) show high dominance and individualism, both in average and in the extreme cases.

$$max(D_T^y) = max(D_T) \wedge max(I_T^y) = max(I_T)$$

TABLE IV
KEY INDICATORS SEGMENTED BY CLUSTER FOUND.

		H_L	H_T	O_L	O_T	P_L	P_T	D_L	D_T	I_L	I_T	C
$n = 115$	$max(X)$	1,0000	0,9032	0,1062	0,0420	0,9722	0,9474	0,5714	0,3871	0,4211	0,3226	0,9605
all	$\overline{X}$	0,4837	0,2083	0,0171	0,0023	0,5998	0,5240	0,2068	0,1319	0,1054	0,0573	0,5289
	$min(X)$	0,0000	0,0000	0,0000	0,0000	0,2174	0,0476	0,0000	0,0000	0,0000	0,0000	0,1728
$n = 32$	$\max(X^r)$	1,000	0,286	0,106	0,010	0,897	0,724	0,571	0,237	0,286	0,176	0,643
Cluster 1	$\overline{X^r}$	0,817	0,053	0,041	0,001	0,610	0,456	0,265	0,137	0,129	0,066	0,480
red	$min(X^r)$	0,483	0,000	0,010	0,000	0,289	0,048	0,080	0,000	0,000	0,000	0,237
$n = 57$	$max(X^y)$	**0,406**	**0,468**	**0,018**	**0,010**	**0,857**	**0,813**	**0,444**	**0,387**	**0,333**	**0,323**	**0,813**
Cluster 2	$\overline{X^y}$	**0,202**	**0,188**	**0,003**	**0,002**	**0,549**	**0,515**	**0,172**	**0,132**	**0,085**	**0,053**	**0,514**
yellow	$min(X^y)$	**0,000**	**0,000**	**0,000**	**0,000**	**0,222**	**0,185**	**0,000**	**0,000**	**0,000**	**0,000**	**0,173**
$n = 26$	$max(X^b)$	1,000	0,903	0,046	0,042	0,972	0,947	0,421	0,316	0,421	0,263	0,961
Cluster 3	$\overline{X^b}$	0,693	0,442	0,017	0,005	0,702	0,628	0,218	0,126	0,129	0,052	0,618
blue	$min(X^b)$	0,469	0,154	0,000	0,000	0,455	0,368	0,000	0,000	0,000	0,000	0,354

Therefore:

Yellow leaders head their sprints very little and there is a clear shift of the initiative towards his/her teams.

D. About Cluster 1 (red)

The right branch of the decision tree in Figure 6 is shared between Clusters 1 (red) and 3 (blue) and in both cases, the heading of the leader is not low $0.406 < H_L^r$. Then, Cluster 1 splits into two alternatives, namely, either the heading of the team is very low ($H_T <= 0.162$) or the team is little collaborative $C <= 0.56$. Let us look at Table IV and Figure 5 to add a little more information.

1) Red leaders head their sprint much longer than yellow or blue leaders together.

$$\overline{H_L^{y,b}} < \overline{H_L^r}$$

2) 50% of red leaders have a heading higher than 80%, that is, they literally head all the sprint.

$$0.8 <= H_L^r$$

3) Red leaders are the most overloaded leaders and they reach the maximum overload level.

$$\overline{O_L^{y,b}} < \overline{O_L^r} \wedge max(O_L^r) = max(O_L)$$

4) Red leaders are the most dominant leaders and they also reach the maximum dominance.

$$\overline{D_L^{y,b}} < \overline{D_L^r} \wedge max(D_L^r) = max(D_L)$$

5) On the contrary, 94% of red teams are under the percentile 15 of heading, that is, red teams barely head their sprint.

$$H_T^r < 0.162$$

6) Red teams and the less collaborative and the most individualistic of all teams

$$\overline{C^r} < \overline{C^{y,b}} \wedge \overline{I_T^r} > \overline{I_T^{y,b}}$$

Therefore,

Red leaders head most part of the sprint in solitaire bearing clearly most of the initiative, and red teams barely share that initiative

E. About Cluster 3 (blue)

All blue leaders do not have a low heading ($H_L^r > 0.406$) and blue teams are the ones with either highest heading ($H_T > 0.343$) or highest collaboration ($C > 0.562$). Let us add, the information in Table IV.

1) Blue leaders show a moderated heading and a moderated overload, but not in excess, it is something in between the yellow and red leaders.

$$\overline{H_L^y} < \overline{H_L^b} < \overline{H_L^r}$$

$$\overline{O_L^y} < \overline{O_L^b} < \overline{O_L^r}$$

2) Blue leaders are more participative than yellow and red leaders together and, furthermore, they reach the maximum level of participation

$$\overline{P_L^{y,r}} < \overline{P_L^b} \wedge max(P_L^b) = max(P_L)$$

3) Despite this, blue teams also show the highest heading, participation, collaboration, and even they reach the maximum level in all these indicators.

$$\overline{H_T^{y,r}} < \overline{H_T^b} \wedge max(H_T^b) = max(H_T)$$

$$\overline{P_T^{y,r}} < \overline{P_T^b} \wedge max(P_T^b) = max(P_T)$$

$$\overline{C^{y,r}} < \overline{C^b} \wedge max(C^b) = max(C)$$

Therefore,

Blue leaders head their sprint, but not in excess, because their team members are also very committed, so the initiative and the effort are shared between the leader and his/her teams in the highest collaborative way

V. At last, the dinosaur

Although it cannot be 100% certain, it is pretty much obvious that this categorization into yellow, red and blue teams and leaders resembles the Kurt-Lewin classification. That is,

- The blue team fits almost exactly as the democratic leader classification. The leader and the team share their effort and show the highest collaboration and participation, without dominance or individualism.
- The red team represents the authoritarian leader, who is very dominant, very overloaded too and does not allow the team to take their own initiative who turn to be more individualistic
- The yellow team represents the laissez-fair leader, that is, a missing leader which produces chaos in the team and the rising of an alternative, unofficial leader who assumes the lost control.

Therefore, it is really surprising to see that, after 7 years of historic data gathered, from nearly 500 students, with different compositions, different duration and different assignments, no matter what you do, there seems to be a clear division into these three categories. In such a way that a simple decision tree, like the one shown in Figure 6 reaches an accuracy of 100%, in only 6 iterations, and its interpretation matches almost perfectly, the theoretical description of these categories.

However, there are some uncertainties about these results which must be highlighted too.

Firstly, **is the activity record R enough to determine all the dynamics within the teams?** Clearly it is not, since there may be conversations, emails, meetings whose relevance with respect to good or bad habits could be high. It is true, but these items are missing. But, undoubtedly, it does reflect the relative effort of the team.

Second, **does a high D_L really mean that the leader might be described as "dominant"?** To be honest: not necessarily, however it is undeniable that when the leader takes part in a story, the greatest effort comes from him/her, so there seem to be actually related.

Or finally, **could a team with a high C, a high P_T and a high H_T and the lowest I_T be considered as "democratic"?** Well it might seem difficult to prove, and furthermore, it is difficult to prove, but again, it is the most likely answer based on the evidences. Indeed, two years ago, leaders of all groups, filled out an well known autoevaluation survey and, guess what? 100% considered himself as a democratic leader.

Obviously, these results are very dependent of the 11 key indicators, if we had selected other indicators instead, the categorization would have been different, it is true, but it is also true that they were selected specifically for that purpose, to highlight these behaviors. Furthermore, it is time to remind the reader that these clusters have been completely shaped by a unsupervised clustering method like R::kmeans() , from a dataset of 118 historic scrum records, with no semantic or biased additions, only the temporal record of activity. The optimal number of clusters has been automatically decided by the known R Packages R::NbClust() and R::clValid(). And that finally, the remaining records have been fed to R::C50, another well known, supervised, classification method, giving a final classification shown in Figure 5. All this **without any human intervention, at all**, and the results ended up being very solid and unexpectedly coherent with the literature. So, this what we have found in the fossils, does look like a dinosaur, doesn't it?

Thus, the main consequence of this work is that it empowers the teacher to get to know these different team dynamics very early, just watching them behave as they want to, and observing their footprints in the historic activity records. There does not seem necessary to make intrusive surveys to the students, just use the rules in Figure 6 as predictors, and, if needed, take corrective measures to achieve the course goals and foster a fair evaluation of the team, avoiding dysfunctions among the students like dominance or abandonment and reinforcing a fair collaboration and participation of all team members . But this is the subject of another study.

References

[1] Guy Brock, Vasyl Pihur, Susmita Datta, Somnath Datta, et al. clvalid, an r package for cluster validation. *Journal of Statistical Software (Brock et al., March 2008)*, 2011.

[2] L. Castillo. The use of scrum for laboratory sessions monitoring and evaluation in a university course enforcing transverse competencies. In IEEE, editor, *International Symposium on Computers in Education (SIIE)*, pages 47–52, 2014.

[3] L. Castillo. Resultados preliminares más significativos tras cuatro años de aplicación de la metodología scrum en las prácticas de laboratorio. In *Actas de las Jornadas sobre Enseñanza Universitaria de la Informática, Vol 2*, pages 47–54, 2017.

[4] Luis Castillo Vidal and Francisca Sánchez Carrascosa. Trabajo en equipo y liderazgo: una propuesta funcional de evaluación de dinámicas internas de grupo. *Actas de las Jornadas sobre Enseñanza Universitaria de la Informática*, 4:327–334, 2019.

[5] Zuzama Covas and Juana Maria. Liderazgo: estilos de liderazgo según kurt lewin y análisis de un caso real. 2017.

[6] Jessica E Dinh, Robert G Lord, William L Gardner, Jeremy D Meuser, Robert C Liden, and Jinyu Hu. Leadership theory and research in the new millennium: Current theoretical trends and changing perspectives. *The Leadership Quarterly*, 25(1):36–62, 2014.

[7] John A Hartigan and Manchek A Wong. A k-means clustering algorithm. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 28(1):100–108, 1979.

[8] Henrik Kniberg. *Scrum and XP from the Trenches*. Lulu.com, 2007.

[9] Max Kuhn, Steve Weston, Nathan Coulter, and R Quinlan. C50: C5. 0 decision trees and rule-based models. *R package version 0.1. 0-21, URL http://CRAN. R-project. org/package C*, 50, 2014.

[10] Kurt Lewin, Ronald Lippitt, and Ralph K White. Patterns of aggressive behavior in experimentally created social climates. *The Journal of social psychology*, 10(2):269–299, 1939.

[11] James MacQueen et al. Some methods for classification and analysis of multivariate observations. In *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, volume 1, pages 281–297. Oakland, CA, USA, 1967.

[12] Barbara Oakley, Richard M Felder, Rebecca Brent, and Imad Elhajj. Turning student groups into effective teams. *Journal of student centered learning*, 2(1):9–34, 2004.

[13] Francisco J Rodríguez-Sedano, Miguel Á Conde, and Camino Fernández-Llamas. Measuring teamwork competence development in a multidisciplinary project based learning environment. In *International Conference on Learning and Collaboration Technologies*, pages 466–479. Springer, 2018.