World Academy of Science, Engineering and Technology
International Journal of Educational and Pedagogical Sciences
Vol:17, No:5, 2023

Empirical Study from Final Exams of Computer Science Courses Demystifying the Notion of 'an Average Software Engineer'

Alex Elentukh

***Abstract*—**The paper is based on data collected from final exams administered during five years teaching the graduate course in software engineering. The visualization instrument with four distinct personas has been used to improve effectiveness of each class. The study offers a plethora of clues toward students' behavioral preferences. Diversity among students (professional background, physical proximity) is too significant to assume a single face of a learner. This is particularly true for a body of on-line graduate students in computer science. Conclusions of the study (each learner is unique and each class is unique) are extrapolated to demystify the notion of an 'average software engineer'. An immediate direction for an educator is to assure a course applies to a wide audience of very different individuals. On another hand, a student should be clear about his/her abilities and preferences - to follow the most effective learning path.

***Keywords*—**K.3.2 computer & information science education, learner profiling, adaptive learning, software engineering

I. Introduction

FINAL Exam by design does not include any new information. It does not introduce new concepts. Final exam covers only those topics that have been covered previously in the course and nothing else. Professors discover something about students and students discover something about themselves. It is all happening very quickly within few hours.

If you walk into a room right before an exam starts, you feel a tremendous charge of energy. When adrenaline kicks in, people learn so much better. In other words, learning does require a jolt, a push into a different level, a concentration that is uncommon in a daily routine.

After all results are in, a congratulatory note is distributed to the whole class and specifically to the student, who achieved the best score. It is never known in advance, who will get the top score. Exam is contributing to thirty percent of a grade. It is administered once, at the very end of a course. Other methods (assignments, projects, discussions) reveal different paths of knowledge. We are not here to debate which path is better. The fact is that each angle is pointedly different. For example, students, who do well on assignments, might do poorly on a final exam. The apparent direction for an instructor is to fully utilize *all* available methods, with the final exam being one of the methods to acquire knowledge.

Alex Elentukh is with the Metropolitan College of Boston University, 1010 Commonwealth Avenue, Boston, MA 02215 USA (phone: 617-335-6135, email: elentukh@bu.edu)

II. Matching the Shape of the Curve with Human Behavior

To start advancing the concept further, we introduce two hypothetical examples with Figs. 1 and Fig2. This enables us to ration about students' behavior in a context of a final exam.

The hypothetical curve on Fig. 1 shows a strong correlation (0.92) between Time and Score. Each red dot reflects performance of one student and the shape of the fitted curve reflects the dynamics of the whole class.

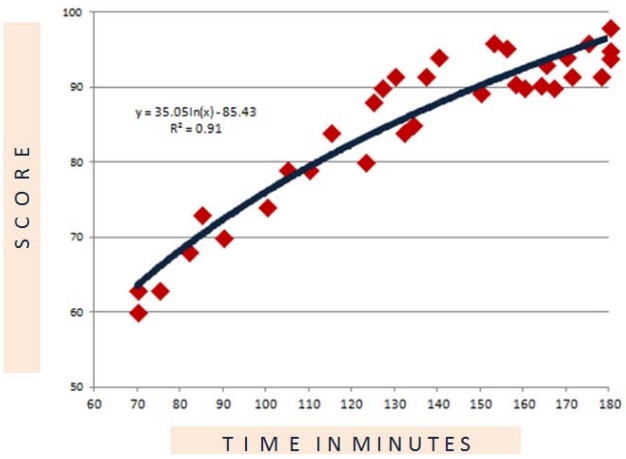

Fig. 1 Positive Correlation (hypothetical data)

As an exercise, we shall interpret the dependence depicted on Fig. 1. Since exam is an open-book, students have the opportunity to look up responses. Those who finished early should have spent more time preparing a fitting response. Those who took the time to carefully examine questions were rewarded with a higher score.

The hypothetical curve on Fig. 2 shows a strong negative correlation (-0.89). Those students, who achieved the highest score, prepared well to the exam. They do not need three hours, since they can cover all questions in a half of the allotted time. Those students, who toiled three hours during exam, were unable to compensate for the missed classes during semester. Hence their grades suffered.

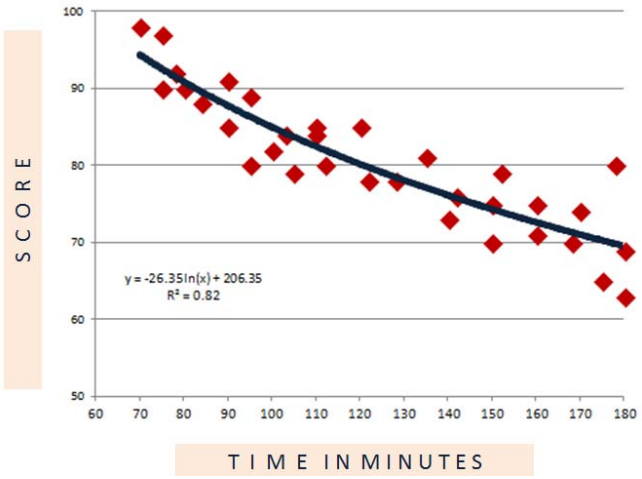

Fig. 2 Negative Correlation (hypothetical data)

The actual chart from a graduate class of 50 students is reflected on Fig. 3. It encompasses both cases from Figs. 1 and 2. Apparently, Time and Score do not correlate. Hence, we need a different method to interpret the collected data - to make it useful and to derive some logical actions.

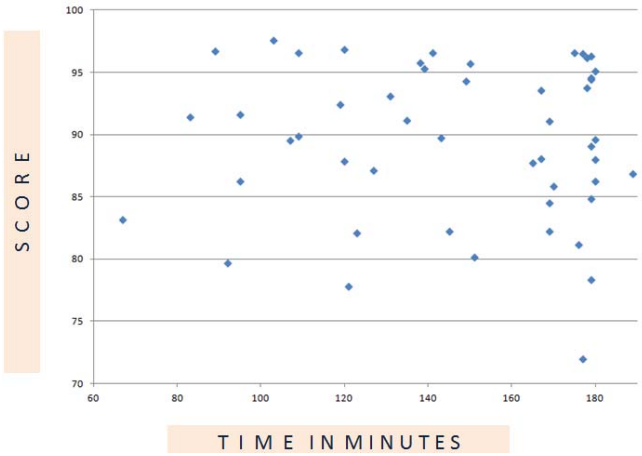

Fig. 3 Final Exam Score VS Completion Time (actual chart)

- Fig. 3 reflects the so-called 'blended class' consisting of both on-line and on-campus students. It appears that the score for final exam of on-campus students is significantly better. The instructor's advice always has been, if it is all possible, to come to the class. On another hand, on-line students do better with a term project. The plausible explanation to this phenomenon is that on-line students usually have an extensive experience in industry.
- A guidance from instructor - to use all time allotted for an exam - has been neglected by a number of students.
- For all five years, as we are running this class, there has been a single person with a perfect score. He scored 100% on all assignments and all quizzes ... for all classes he ever took at university. He grew up in India and he is currently working as a Software Lead Engineer. Having said this, to get an 'A' for this class, one does not need to score a hundred percent.

III. Four Personas

Diversity among students is too significant to assume a single face of a learner. An aggregation of individual virtues is bound to distort the overall model and render it useless. In this study, we use four personas. Fig. 4 divides the whole space between Score and Time into four corresponding areas.

Fig. 4 Four Personas

Further we provide a brief characterization to each quadrant.

- *Persona (1) Genius:* These folks do things that are clearly above and beyond. They can be characterized as having exceptional innate capabilities, coupled with a relentless drive.
- *Persona (2) Expected Behavior:* Most students are in this category. It is a safe ground known for its logical and prudent actions.
- *Persona (3) Over Confident:* An experienced software engineer, who does his 'thing' and sees little value in broad topics. The classic Dunning-Krueger curve reflects this behavioral pattern.
- *Persona (4) Need External Help:* A foreign student who is brought up in a very different environment and is unable to keep up with the course.

Definition of personas is a prudent first step of any software project. Prior to building a system, it is important to identify several target archetypes by segmenting the whole user space. In our context, students (as users) are represented with four aggregate personas. With a large class (of over 50 students), red dots appear in each of the four areas. For example, if most dots congregate in the bottom-right corner, it is indicative that the whole course (not just the final exam) has to be revisited.

For comparison, Mojarad et al. [1] describe a study of creating 5 clusters of students based on 6 academic and behavior characteristics. Such clusters are assembled during the earlier part of a course, so then to be used at the later part.

IV. Manifestations of Over-Confidence

In many cases, participating in a discussion about overconfidence is similar to explaining to a color-blind person the intricacies of a flashy user interface. The taxonomy of such conversation is quite limited.

The only way to meaningfully diagnose this catastrophic cognitive bias is as follows, "If you find yourself at the bottom left corner of the chart, it means you are overconfident". Reason [2] covers an extensive catalog of software deficiencies with 'over-confidence' being one of the biggest buckets that is most commonly full. Brenner [3] and Hilary and Hsu [4] have an in-depth analysis of various biases of being 'over-confident'.

V. An Instrument of Collaboration

For this study, to preserve confidentiality, names of students and their classes were redacted. At the beginning of a class, each student publishes a personal profile, to establish a connection with other students.

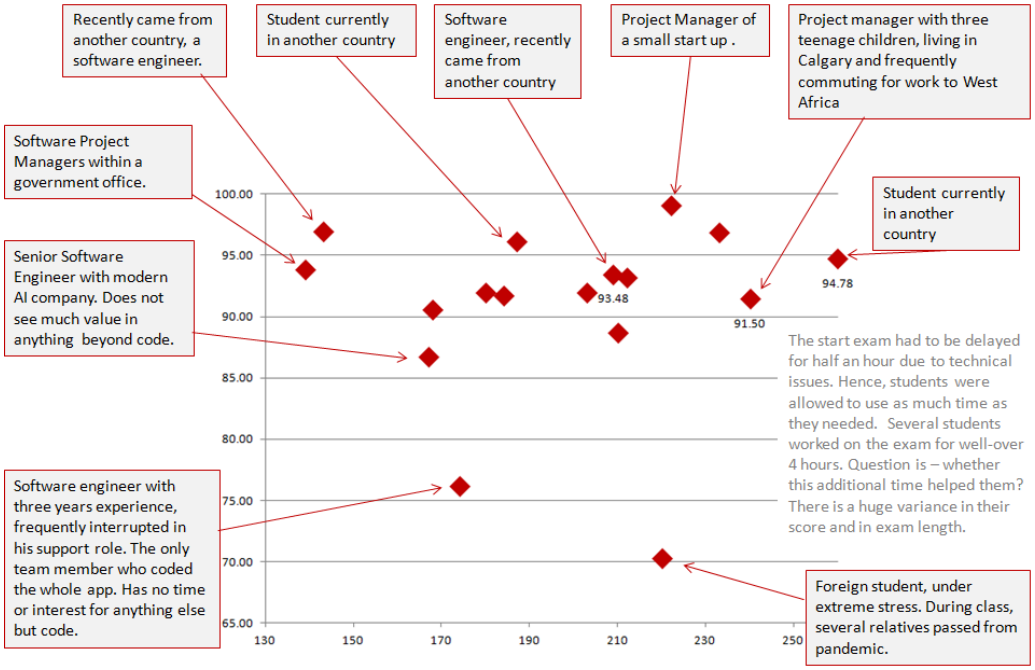

Fig. 5 Score VS Time (actual example)

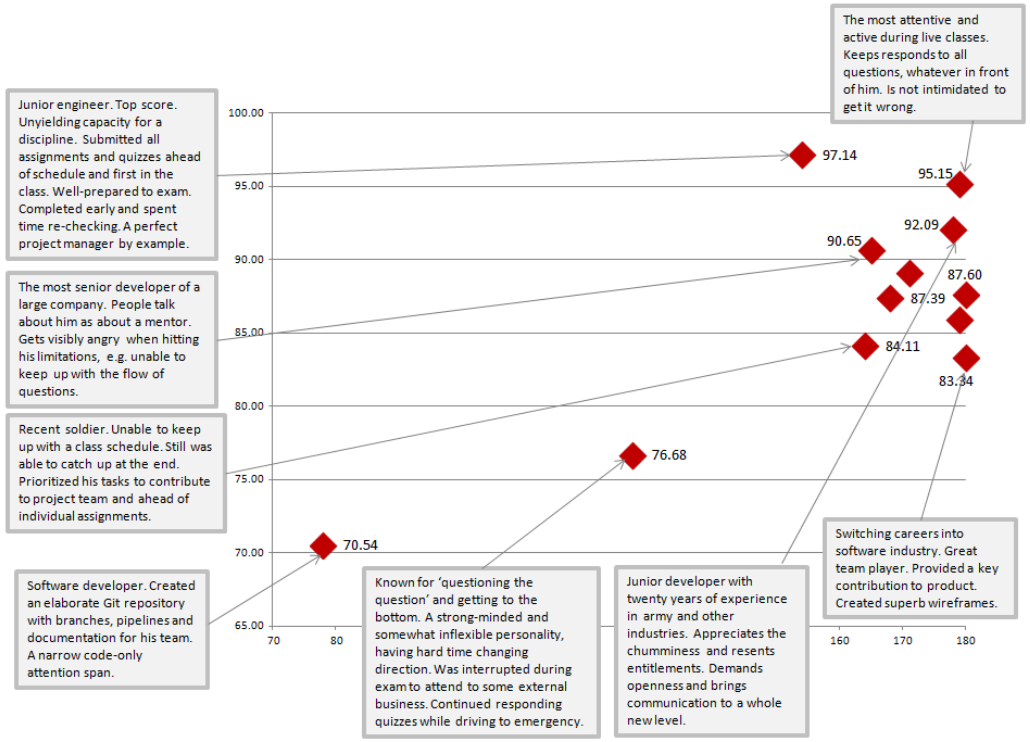

Fig. 6 Score VS Time (actual example)

When a learner finds his/her unique dot on a class chart, it serves as a strong motivation. Such chart is distributed and discussed within each class. It grows into an instrument of collaboration to brainstorm and ration about,

- why so many dots congregate in a certain area, or
- why several folks left exam so abruptly, etc.

Looking at these real examples of charts, Figs. 5-7, it is logical to map a student's performance into one of four predefined personas.

A far-reaching conclusion is derived from the class where the final exam was inadvertently extended due to technical issues with blackboard. Students were offered to spend on exam as much time as they wanted. Instead of usual three hours, exam was run for well over four hours. Apparently, this additional time did not improve grades. Looking at Fig. 5, one can still observe a huge variance among scores. The innate capabilities and preparedness to exam were the main drivers for a score.

- A developer with years of experience in industry is receiving a low score of 75, after spending just 175 minutes on the exam. Such dynamic belongs to persona # 3. The student falsely assumes that he knows everything there is to know about the topic.
- A student, who is currently in another country, with both parents passed from pandemic, is scoring just below 70, while exhausting most of the allotted time. This belongs to persona #4. Understandably, he has other things on his mind.

Dynamic of a relatively-small and well-jelled class is reflected on Fig. 6. One can see a familiar picture, as the dot of an experienced engineer is positioned at a bottom left corner. The best score belongs to the person with an unyielding capacity for a discipline. During semester, all assignments and quizzes were submitted ahead of expected milestones. This behavioral characteristic is translated into the top grade within the whole class.

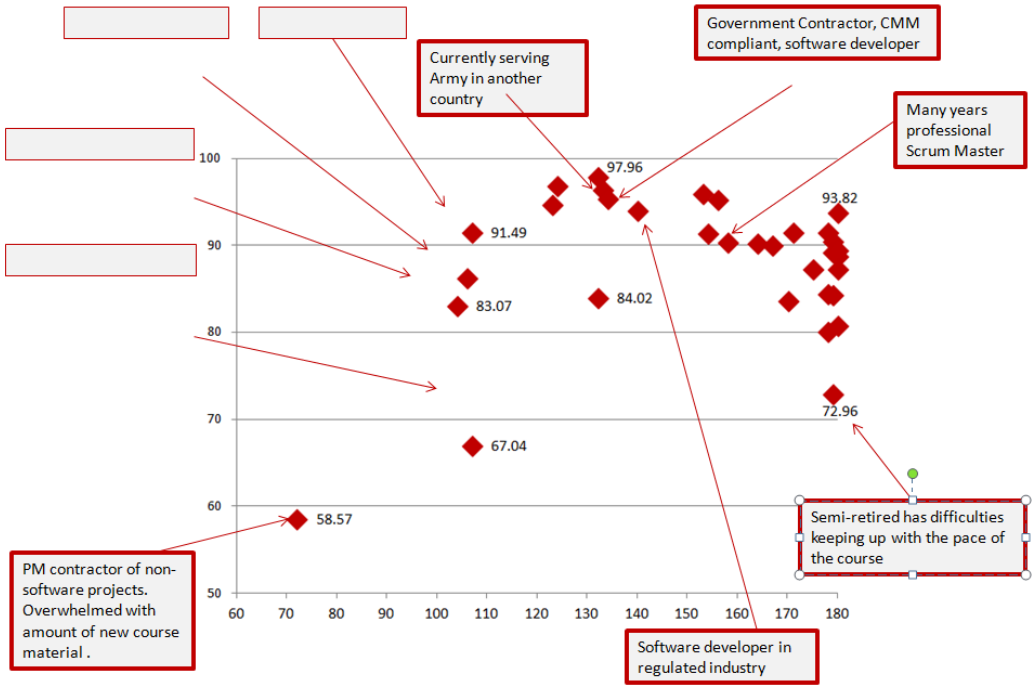

Fig. 7 Score VS Time (actual example)

Diversity of backgrounds is apparent from these actual examples, Figs. 5-7. There is a student serving in Army in Korea, semi-retired IT support staff in Vermont, and then, there is a project manager living in Calgary, while commuting to work to West Africa. What unites all of them is the on-line CS class, where participants quickly grow into a congruent team, with the only goal to learn.

Dynamic of a small class with nine students is reflected on Fig. 8. One can easily recognize two well-defined groups. First group submitted exam at about two hours mark. The second group worked hard to exhaust all allotted time and submitted at three hours. There is an apparent anchoring effect, as folks within each group do influence each other. We can talk about a certain group culture when people within each group discuss their issues and sit together during lectures. One person at the Group A is a senior software engineer, who sets the behavior pattern. Apparently, he is not interested in absorbing the complete material of the course and is happy with a limited scope of whatever directly pertains to his work. This is reflected in his low score. The top student of the class belongs to the group at the right. His innate technical talents are combined with a great capacity for a discipline. He paid his own money for tuition and subconsciously needs to justify this investment. It would be uncharacteristic of anyone from this group to submit exam early.

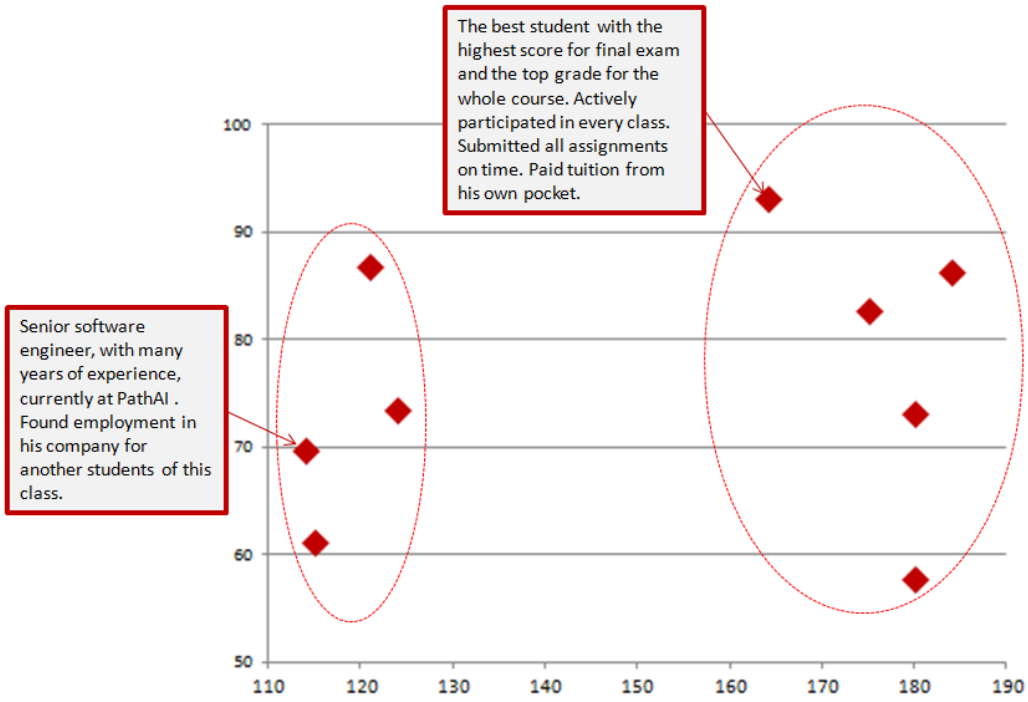

Fig. 8 Score VS Time (actual example)

VI. Prudent Empirical Research

The traditional challenge of an empirical research in teaching is to find a parameter (or a set of parameters) that has a strong correlation with an outcome (final grade). In a situation of a shortage of a magic parameter, we look at what we have instead.

The indisputably-exact parameter in front of us is 'time spent doing final exam'. None would dispute the precision of this parameter. It is retained forever for any faculty to observe in a Learning Management System, in our case, in Blackboard. Most importantly, it orchestrates the context for any possible research activities.

It seems unfortunate that parameters, 'score' and 'time spent during final exam' are *lagging* indicators. They reflect on what has already happened. It would be more useful, if we are able to leverage a *leading* indicator to improve the learning. Still, considering that to obtain a graduate degree, a student has to go through a dozen courses. A learner can definitely benefit early in a program from becoming aware of his/her profile.

VII. The Importance of Being a Fast Learner

The subconscious questions remain - whether being a quick learner is important; or the quickness is just a yardstick, the starting point? How much effort should we invest in being quick; or we do not need this virtue at all? What happens if we are not fast, does it make a difference?

In any case, if you can respond to these questions quickly, it is definitely a good sign, it is a manifestation of some other useful trait. *Developing quickness is bound to draw out other virtues, which is exactly what the final exam does.*

During a live session, instructor uses all possible methods to encourage students to be active. An instructor says, "...at this point, there is no right or wrong answer; what counts is your participation". Some students do it seamlessly; they respond immediately to make sure they are, in fact, a part of a continuous conversation. Their 'pipeline between short and long memory' is squeaky clean, which enables them to chime in, or to bring about a supporting fact from personal experience, or to fetch some alternative interpretation. Such virtue of 'thinking on your feet' is most constructive in a back-and-forth, collaborative dynamic, [5], which is opposite to a command-and-control mono-directional and infinitely less effective dynamic.

VIII. Conclusions

Learner profiling occupies a large part of today's research in teaching. The flow of published papers has increased with the intense proliferation and dominance of on-line courses. Most papers present various interesting theories. Still, the application of research remains scarce and far in between [6]-[8].

Our paper summarizes an empirical study with its conclusions put into practice immediately. With reference to Fig. 4 'Four Personas', it should be noted that moving your dot from bottom to top and from right to left is a significant effort. In most cases such effort extends beyond the scope of a university course. Still, becoming aware and accepting your talents and weaknesses is a sensible first step.

The other side of being over-confident is to think that you do not know, when in fact you do. In this case, a little push will result in a pleasantly surprising improvement.

An educator has to assure that a course has a certain depth. As the same message applies to very different individuals. To this end, a term project [9], [10] is an effective method that enables members of a team to learn from each other. Small teams are composed with folks of complementary backgrounds. One selects a role that is most familiar, to be able to impart the

skills and knowledge to the rest of the team. This is opposite to selecting a role that is least familiar, while focusing on personal interests.

Responding a question, "who is the target (customer, persona) of a course?" An educator should confirm that *everyone* who came to the class will benefit from the class.

In our paper, learning habits and preferences are extrapolated from dynamics of a final exam toward the universe of acquiring general knowledge. Any collaboration on these research topics is welcomed.

REFERENCES

[1] Shirin Mojarad, Alfred Essa, Shahin Mojarad, Ryan S. Baker, "Data-Driven Learner Profiling Based on Clustering Student Behaviors: Learning Consistency, Pace and Effort". https://link.springer.com/chapter/10.1007/978-3-319-91464-0_13
[2] James Reason, "Human Error", Cambridge University Press, 1990.
[3] Point Lookout, Chaco Canyon Consulting, February 2, 2022, Volume 22, Issue 4.
[4] Gilles Hilary and Charles Hsu. "Endogenous overconfidence in managerial forecasts." *Journal of Accounting and Economics* 51:3, (2011), 300-313. Retrieved 17 January 2022.
[5] *Susanna Barrineau,* "Students as Change Agents – Reorienting Higher Education Pedagogy for Wicked Times", Uppsala University.
[6] Sallie Gordon, "Systematic Training Program Design", Prentice Hall, ISBN-13: 978-0131003897.
[7] Essaid El Bachari et al. "An Adaptive Learning Model Using Learner's Preference", 2010, Cadi Ayyad University.
[8] Jose Manuel Marquez Vazquez, et. al. " Designing adaptive learning itineraries using features modeling and swarm intelligence". Springer-Verlag London Limited 2011.
[9] Alex Elentukh, Vijay Kanabar, "Improving Teaching and Learning Effectiveness of Computer Science Courses – Case Study ", CSECS 2019.
[10] Sally Fincher, Marian Petre, "Project-Based Learning Practices in Computer Science Education". December 1998.