Automated Theme Allotment to Optimise Learning Outcomes in Robotic Competition

Gayatri Belapurkar
Department of Information Technology
Vivekanand Education Society's Institute of Technology
Mumbai, India
gayatri.belapurkar5@gmail.com

Aniruddha Chauhan
Computer Science Department
Bharati Vidyapeeth's College of Engineering
New Delhi, India
aniruddhac7@gmail.com

Aditya Panwar
Computer Science & Engineering Department
Indian Institute of Technology Bombay
Mumbai, India
adityapanwar01@gmail.com

Vikrant Fernandes
Computer Science & Engineering Department
Indian Institute of Technology Bombay
Mumbai, India
vikrant.ferns@gmail.com

Kavi Arya
Computer Science & Engineering Department
Indian Institute of Technology Bombay
Mumbai, India
kavi@cse.iitb.ac.in

***Abstract*—In the e-Yantra Robotics Competition we teach robotics through a competition paradigm. Here, 'Themes' or problem statements, are allotted to participants who clear a preliminary selection test. Theme allotment to student teams is a challenge and is time-consuming since we need to fit a Theme to a team capable of addressing the challenges denoted by a Theme. This allotment of Themes considers multiple factors such as team members' experience, expertise, Theme complexities, domain requirements and other constraints set on a Theme by the respective Theme designers. Various machine learning techniques were analyzed and experimented with; however, the problem of project tagging could not be solved by machine learning algorithms alone due to certain randomness in data and human bias involved in previous Theme allotment techniques. We thus devised a new metric that projected multiple parameters onto three dimensions, namely Algorithm, Mechanical and Electronics. Further clustering and analysis were done based on these three dimensions. The predicted clusters were subjected to conditional random sampling based on Theme constraints for actual project allocation to ensure an unbiased and a better allotment of Themes. This led to allotment of Themes such that it provided control to Theme designers, whilst automating the time-consuming process.**

Keywords—Theme allotment, project-based learning (PBL), Clustering, Sampling, metric fabrication, e-Yantra Robotics Competition

I. Introduction

e-Yantra, a project at Indian Institute of Technology Bombay, was introduced to develop ground-breaking digital pedagogies to deploy and enhance the concepts of robotics and embedded systems in order to train both college faculties and students across the whole country. e-Yantra does so by having various initiatives under its belt such as:

- e-Yantra Lab Setup Initiative (eLSI) under which fully equipped robotic labs are established in the colleges across the country. 361 labs have been setup under eLSI as of writing the manuscript.
- e-Yantra Ideas Competition (eYIC) where a student team along with a faculty mentor from an eLSI college proposes an idea and implement it after relevant tweaks and feedbacks.
- e-Yantra Resource Development Center (eYRDC) that provides exclusive content to eLSI labs that can be used to teach more effectively. It also provides a platform to various faculties to show their teaching prowess in their respective domains.
- e-Yantra Farm Setup Initiative (eFSI) that enables agricultural testbed over IoT technology for automation.

One way e-Yantra indulges the students in the concepts of embedded systems is through an annual e-Yantra Robotics Competition (eYRC) which aims at inculcating knowledge in robotics, embedded systems, image processing, machine learning and micro-controller programming. The competition aims at teaching these core concepts through a hands-on practical learning-based system as opposed to the rote theoretical learnings of the traditional classroom system. The students participate as a team of four student members that have the flexibility of being from any year and studying in any branch. Participating teams use these concepts along with previous know-how to solve abstracted real-world problems [1]. These gamified problem statements or Themes are abstracted on a static arena comprising of numerous static and dynamic components which teams must devise a solution for. Since its inception in 2012 (eYRC-2012), this competition has been growing exponentially while showing that the students learn better while competing. Each version of the competition comprises of various Themes under a wide umbrella or domain viz. Agriculture, Urban Planning, Warehouse Management, Space Exploration and Forest Safari. All versions of eYRC are completely conducted online. This competition is conducted in multiple stages, each stage comprises of tasks related to respective Theme; with each stage eliminating several teams. All these tasks go hand-in-hand with constant feedbacks and guidance of the Theme designers via various online discussion forums viz. Piazza discussion forum [2], Google hangouts etc. The first stage consists of a preliminary selection test, which tests basic Aptitude, Programming skills and Electronics knowledge of the students. This test contains questions of varying complexity levels and the complexity level for questions is regularly revised using an auto-tagging algorithm [3], after which a set of students are selected for the next stage

We thank Ministry of Human Resource Development (MHRD), Govt. of India for sponsoring e-Yantra Project under National Mission on Education through Information and Communication Technology (NMEICT).

of the competition. These students are then assigned a Theme, based on the background of its members [4], that they have to work on and implement over a course of six months. The teams that perform well in this stage qualify for the next stage which includes the actual hardware implementation of the allotted Theme. All the necessary hardware required for any Theme is shipped by e-Yantra without any charge to the participating teams. Among these teams, the top five to seven top scoring student teams who successfully implement the solution of the Theme are selected for the grand finale that are held at Indian Institute of Technology Bombay.

The latest version of eYRC i.e. eYRC-2018 had 7244 teams that registered for the competition that translates to 28,976 participants, of which 5790 teams appeared for the selection test. eYRC-2018 comprised of seven different Themes under a wider domain of Forest Safari. With this tremendous increase in the number of participating teams, the process of allotting these Themes manually got time-consuming and cumbersome.

In order to reduce the manual effort and to introduce a few more factors while allotting Themes, we devised a new system to automate this process. This system essentially reduces various factors that includes the prior eYRC experience of the participants (if applicable), field of expertise of the students in a team based on their course/department they are studying in, the year they are studying in and the team's score distribution in Aptitude, Programming and Electronics in the preliminary selection test, down to an 'Algorithm' component, a 'Mechanical' component and an 'Electronic' component, which is further used to process the data and allot Themes to competition participants.

For further allotment of Themes, various learning approaches such as ensemble models, clustering methods and probabilistic machine learning approaches were tried and tested. However, this did not provide an objective measure to the allotment of Themes. Thus, a different mechanism was designed to overcome the drawbacks of manual tagging, to reduce the time, and to enable unbiased and more effective allocation of Themes.

II. Motivation and Related Work

Theme allocation or 'Project Tagging' in eYRC has been done manually over the years, based on a total average score of the four students in a team obtained in a selection test, the departments that the students come from, previous years' eYRC experience, specifications of the respective Themes , and whether the team belongs to an e-Yantra Lab Setup Initiative (eLSI) [5] college.The main motivation of this work was to automate this process of project tagging or Theme allotment in order to reduce human effort and time along with making it a fair and more efficient process free of human bias. It is estimated that the additional factors considered in the new system will lead to better allocation of Themes, reduce bias, and is expected to bring an increased success in terms of full hardware implementation of the problem statements i.e. Themes assigned to competing teams in the competition.

Similar work was done for dynamic profile classification and project allotment that used an ensemble based learning model for classification [6] and focused on automating the process of project allotment, while increasing the efficiency.

III. Analysis of Current Method

A. *Themes*

Each Theme that is introduced in eYRC has a theme designer who has implemented the Theme before it is finalized and ready to be deployed for eYRC. The theme designer has experience and expertise in the technology stack required for a particular Theme and has a good idea of the level of complexity of the Theme and also about issues or obstacles that students may face while implementing the Theme and what may be required to overcome these issues. Thus, each Theme designer has his/her own criteria and specifications as to which teams should be allotted that Theme, in order to ensure that the teams are diligently able to implement the theme and for proper distribution of teams in the competition. Moreover, this distribution also reduces the feedback and guidance time of the designers

A 'Theme complexity matrix' is hence fabricated for this purpose by theme designer that gives the 'Algorithm', 'Mechanical' and 'Electronics' complexity of that Theme in terms of the percentage and the Theme designer's preference of students on the basis of their departments.

B. *Insights from the current allotment method*

The current method of allocation of manual tagging was analyzed from the years 2016 to 2018. What was observed during the analysis was that in some cases there was an iota of bias, some inherent randomness and a few discrepancies in the allocation of projects. This was because there are a number of factors taken into consideration when allotting a project and the sheer scale of teams participating in the competition as human bias while allotment gives rise randomness in the data. For instance, a Theme that required a team to have higher knowledge of a particular dimension, a few teams that were actually allotted that Theme may have come from the section that scored comparatively lower in the selection test. Similarly, for a comparatively easier Theme, teams that performed better on the selection test might get included at times. Thus, very little inferences could be made from the current allotment method due to the extent of manual involvement and human errors.

The plot shown in Fig. 1 depicts the randomness introduced in Theme allotment due to human intervention and lack of an ordered system. Each colour represents the teams that were allotted a particular Theme, which is plotted along the respective teams' Algorithm, Mechanical and Electronics scores. As can be noticed, a Theme does not appear to be allotted as per the teams' calibre.

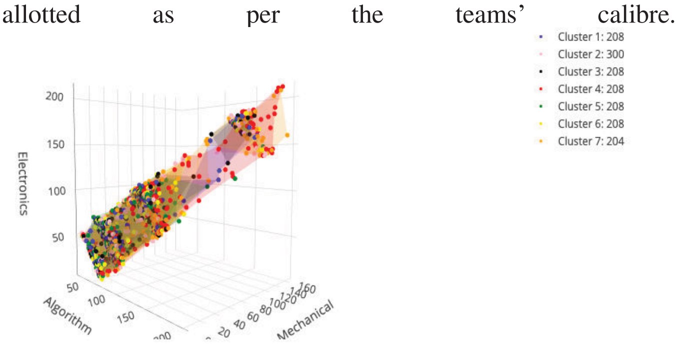

Fig. 1. Plot of Theme allotment in eYRC-2018 as per the team's 'Algorithm', 'Mechanical' and 'Electronics' component

In order to get an insight as to how well this system of manual tagging works, results of the grand finale of the competition (Fig. 2) were analyzed for the years 2016-2018. It was noticed that although a Theme was comparatively harder to implement as devised by the Theme designers, the top five to seven finalists selected for that Theme may have scored average or comparatively lower marks in the selection test in comparison to teams that were not selected as finalists. This suggests that Project-Based Learning (PBL) in eYRC provides a functional upgrade to the technological knowledge of students and that simply achieving a higher score in the selection test alone does not necessarily translate to good performance in the practical implementation of robotics tasks.

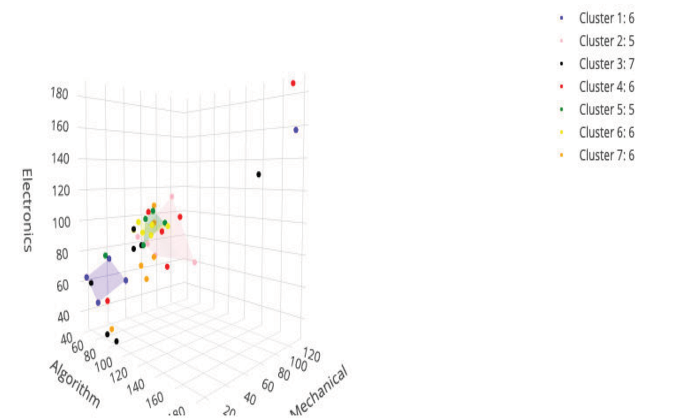

Fig. 2. Plot of in 2018 winners in each Theme as per the team's 'Algorithm', 'Mechanical' and 'Electronics' component

IV. Approaches Considered

We considered machine learning algorithms to make the process of Theme allotment easier, faster and as close to manual tagging as possible. We started off with unsupervised learning methods such as K-means [7] algorithm. However, learning without any labelled data would be problematic and hence not give the desired results in this case, as manual tagging was done in a different way, with multiple other constraints that the clustering model did not consider. That led us to semi-supervised learning as the next alternative. Decision tree algorithms [8] [9], random decision forests algorithm [10], k-Nearest Neighbors (kNN) [11], support vector machines [12] [13] and ensemble methods [14] such as XGBoost were tested, however, the results obtained were not good as the maximum accuracy achieved was 20%. This was mainly due to the randomness in the data, which could not be deciphered by the models. Thereafter, we tried Probabilistic machine learning approaches such as Bayesian Networks. This too, was not an efficient method as it gave a score of -0.644 when compared against manual tagging. It was also time consuming and required hand-picked training data, which further did not translate to a good accuracy while allotting Themes. After trying out multiple approaches, it was concluded that due to the randomness in data and the high amount of human bias involved while allotting a theme, project tagging was not achievable using machine learning models alone and an entirely new system for Theme allotment was required. Thus, a system was designed which would consider all the relevant factors for Theme allocation and overcome the drawbacks of manual tagging.

V. Devised Method

A. Factors considered

The proposed method takes into account additional factors/constraints as opposed to the current manual tagging method in order to make project tagging more efficient and free from human bias. The additional constraints considered in the said method are as follows:

1) Year of Study: The year of study of a student reflects the amount of experience and knowledge that the student has in his/her field of study. A third year Bachelor of Technology (B.Tech.) student in Computer Engineering will on an average have more knowledge than a second year Bachelor of Technology student Computer Engineering student.

2) Department: Departments include the various different departments that students study in, such as Electronics Engineering, Mechanical Engineering, Computer Engineering, Electrical Engineering, Computer Science, Information Technology, Electronics and Telecommunication Engineering and so on.

3) Prior eYRC experience: If any team has one or more members who have participated in the previous version of the competition, they have an added advantage. The stage and certification level upto which these student went in their prior experience is an important factor that may add to the advantage. A student eliminated in the second stage will have lesser advantage as compared to a student who was a finalist in any of the previous years' Themes.

4) Test Scores per section: Along with the overall score obtained by the team in the test, it was deemed necessary to look at the individual distribution of the team's scores so as to identify the strengths and weaknesses of a team. For instance, consider a team that has a high score and consists of final year B.E. students of Computer Department. Their strength is identified to be 'Algorithm' in the selection test. If this team is given a mechanics intensive Theme to implement, they may not be able to give their best, while on the other hand if they are given an algorithm intensive Theme, they can make the best use of their strengths and it can play to their advantage in order to implement the Theme most effectively.

B. Metric Fabrication

We came up with a metric that would combine all the factors and reduce them down to three major dimensions: 'Algorithm', 'Mechanical' and 'Electronics', along which all further processing would take place. For executing this, each of the factors were given weights along the three dimensions and their sum gave the total 'Algorithm', 'Mechanical' and 'Electronics' scores. The various sections were reduced as follows:

1) Year that a student is studying in: The year that a student is studying in is given a weight according to seniority. This weight is then equally added in all the three dimensions of the metric.

TABLE I. WEIGHT ASSIGNMENT ACCORDING TO YEAR THAT A STUDENT IS STUDYING IN

Year Studying In:	Weight Assigned
1st Year	1
2nd Year	2
3rd Year	4
4th Year	8

2) Department: The Departments that students come from are compressed into three 'Branches', namely Computer, Electronics and Mechanical. For example, students from Computer Engineering, Computer Science, Information Technology etc. are all classified into 'Computers' whereas students from Electronics and Communication Engineering, Electrical Engineering, Electrical and Electronics Engneering fall into 'Electronics' Branch. Based on the Branch that the student is classified into, we give that student weights in each of the dimensions of the metric.

TABLE II. WEIGHT ASSIGNMENT ACCORDING TO DEPARTMENT

Year Studying In:	Algorithm Weight	Mechanical Weight	Electronics Weight
Computers	8	1	2
Mechanical	2	8	2
Electronics	4	2	8

3) Prior eYRC experience: The prior eYRC experience that a student has is given weights based on the stage upto which the student lasted in the competition.

If a student went upto Certification Level 1, who are the selected few and compete in the grand finale, are given the highest weight. Students who completed the Theme implementation but were not able to secure the top five to seven spots are given the second highest weight of 4 corresponding to Certification Level 2. On the other hand, students who partially completed the Theme solution fall in the Certification Level 3 category, while those who qualified only for Certification Level 4 - cleared only the first stage, are given the least weight. If a student does not have any prior eYRC experience, then no weight is added from this section. This weight is equally added along all the three dimensions of the metric.

TABLE III. WEIGHT ASSIGNMENT ACCORDING TO PRIOR EYRC EXPERIENCE

Certification Level:	Weight Assigned
Certificaion Level 4	1
Certificaion Level 3	2
Certificaion Level 2	4
Certificaion Level 1	8

4) Test Scores per section: The selection test that was conducted as a preliminary stage measures the student's performance in Aptitude, Electronics and Programming [3]. The Programming score is added to the 'Algorithm' component in the metric, the Electronics score is added to the 'Electronics' component and the 'Aptitude' score is divided by three and added to all the three components of the metric equally.

C. Clustering

The teams were then clustered on the basis of the calculated metric using k-means clustering [7]. This was done as an ideal scenario since eYRC-2018 has seven distinct themes and hence should translate to seven different clusters as shown in Fig. 3. However, in such case every Theme has its own intrinsic value with different Algorthm, Mechanical and Electronics score. Also, two or more Themes might have overlapping requirements due to which, increase in the clusters will increase in biasness. Since we are k-means, every cluster will have different number of teams which is not a fair distribution of the Themes. So, we came up with a better way by dividing the clusters into three parts as plotted in Fig. 4. The clusters give teams with a 'low' score on the metric, an 'average' score and a 'high' score. The reason for using a clustering approach is that it helps us make clusters of homogeneous data-points, which in our case are the metric scores of teams in the three axes.

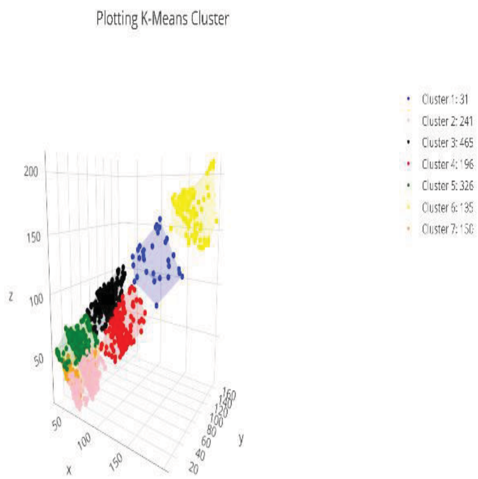

Fig. 3. Plot of teams in eYRC-2018 participants divided into seven clusters as per the number of Themes

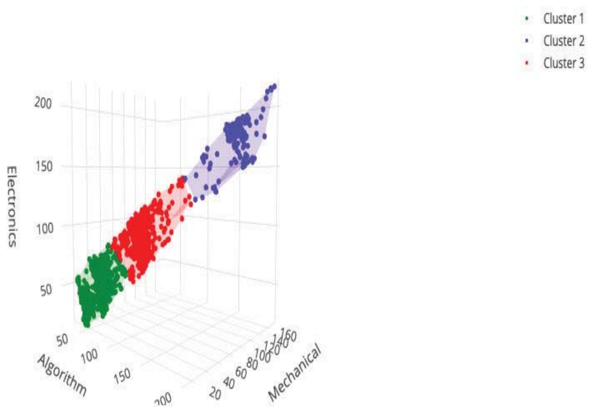

Fig. 4. Plot of teams in eYRC-2018 participants divided into three clusters as per the team's 'Algorithm', 'Mechanical' and 'Electronics' component

Clustering enables us to get a cluster of similar teams based on our metric, which will reduce the randomness and bias when we sample from a cluster because of the similarity in a cluster.

D. Sampling

Once the three clusters are generated, an entire cluster report is given to the Theme designers. Thereafter, they decide the parameters along which they want teams in their Theme. The teams in eYRC that come from eLSI colleges are considered separately as it aids in achieving the objective of the e-Yantra Lab Setup Initiative which is to create an ecosystem where students in engineering colleges from across the country can learn through PBL [5] [15] [16]. Therefore, the parameters considered while sampling teams are essentially the number of teams they want from each of the three clusters, the cut-off scores for their Theme along the three dimensions, the number of eLSI teams from each cluster and the minimum number (if any) of students from each branch. For example, a Theme may require that there be 10 teams from the 'Low' cluster, 20 from the 'Average' cluster and 30 from the 'High' metric cluster, 4 teams in all the clusters from eLSI colleges and at least one person in the team should have a mechanical background for the implementation of the Theme. These conditions are taken as input in a .csv file.

Thereafter, teams are randomly sampled according to these conditions and allotted themes. In the end, a .csv file is generated, which maps teams to their allotted Themes. In case if teams are remaining and are not allotted any Theme they are returned in the same file with no Theme allotted and are left to the discretion of respective Theme designers.

VI. Results

The system, on implementation, automates the process of theme allotment and saves considerable human effort and time. The total time taken for project allotment came down to not more than 6-10 minutes. It not only reduces the manual effort behind project allocation but also provides a more elegant solution for proper allocation of the projects, in our case Themes. In this method, the only human intervention required is for Theme designers to specify the constraints that the teams being allotted a Theme should satisfy so that they are able to do what the Theme demands of them.

VII. Conclusion

We have demonstrated that there is a considerable gain in the knowledge of students and improvement in their performance due to project based learning imparted in eYRC and shown that scores obtained by students in the selection test based on theoretical knowledge does not necessarily translate to successful practical implementation of competition 'Themes'. Also, as observed from the study and the various supporting figures, it is better and more intuitive to use a k-means cluster of three based on the low, average and high scores rather than a fine-grained set of more clusters which will introduce biasness that increase with increase in the clusters due to sampling. Theme designers can then use these clusters to assign the Themes selecting a chunk of student teams from each of these clusters which proves to be a fair distribution among all the Themes. The additional constraints that have been considered in the allotment of 'Themes' are expected to play a role in effective and efficient allotment of 'Themes'.

The new process promises to save manual effort and time and will make theme allotment free of human bias as teams are randomly sampled based on conditions set by respective Theme designers. This is expected to bring about an increase in the number of successful practical implementations of the Theme as the students can play to their strengths which in turn will familiarize the participating students with latest technologies in their related domain while performing on a practical platform and observe the results. After all, the e-Yantra Robotics Competition is really a robotics MOOC designed as a competition. We need to ensure effective training to the maximum number of participants to fulfil our goal of creating effective ways to inculcate core concepts in robotics and mold the students into practically oriented engineers.

Acknowledgment

The e-Yantra Robotics Competition is administered and managed by the e-Yantra staff at Indian Institute of Technology Bombay (IIT Bombay). We acknowledge the all the contributions and support of e-Yantra project staff especially, Omkar Manjrekar in providing his valuable feedback in the initial draft. We would also like to take this opportunity to thank all the colleagues here at Indian Institute of Bombay that provided valuable inputs in the nascent stage of the study. We are also extremely grateful to Ministry of Human Resource Development (MHRD), Government of India, for funding the e-Yantra project under the National Mission for Education using Information and Communication Technology (NMEICT).

References

[1] e Yantra, "e-Yantra," http://www.e-yantra.org/.

[2] M. Vellukunnel, P. Buffum, K. E. Boyer, J. Forbes, S. Heckman, and K. Mayer-Patel, "Deconstructing the discussion forum: Student questions and computer science learning," in Proceedings of the 2017 ACM SIGCSE Technical Symposium on Computer Science Education, SIGCSE '17, pages 603–608, New York, NY, USA, 2017. ACM.

[3] S. Krithivasan, S. Gupta, S. Shandilya, K. Arya, and K. Lala, "Autotagging for massive online selection tests: Machine learning to the rescue," in 2016 IEEE Eighth International Conference on Technology for Education (T4E). IEEE, 2016, pp. 204–207.

[4] S. Battur, M. S. Patil, P. Desai, M. Vijayalakshmi, M. M. Raikar, P. Hegde, and G. H. Joshi, "Enhancing the students project with team based learning approach: A case study," in 2016 IEEE 4th International Conference on MOOCs, Innovation and Technology in Education (MITE), Dec 2016, pp. 275–280.

[5] S. Krithivasan, K. Lala, K. Arya, S. Shandilya, S. Jain, P. Manavar, and S. Patii, "e-yantra lab setup initiative: Sustainable knowledge creation and scalable infrastructure creation at engineering colleges," in 2014 IEEE Frontiers in Education Conference (FIE) Proceedings. IEEE, 2014, pp. 1–8.

[6] S. Tangadle Gopalakrishna and V. Varadharajan, "Ensemble learning based voting model for dynamic profile classification and project allotment," 11 2018, pp. 77–88.

[7] A. C. M. . S. Guido, "Introduction to machine learning with python: A guide for data scientists," 09 2016, pp. 168–181.

[8] A. Gron, "Hands-on machine learning with scikit-learn and tensorflow," 06 2017, pp. 215–231.

[9] A. C. M. . S. Guido, "Introduction to machine learning with python: A guide for data scientists," 09 2016, pp. 70–83.

[10] A. Gron, "Hands-on machine learning with scikit-learn and tensorflow," 06 2017, pp. 232–258.

[11] A. C. M. . S. Guido, "Introduction to machine learning with python: A guide for data scientists," 09 2016, pp. 35–44.

[12] A. Gron, "Hands-on machine learning with scikit-learn and tensorflow," 06 2017, pp. 191–214.

[13] A. C. M. . S. Guido, "Introduction to machine learning with python: A guide for data scientists," 09 2016, pp. 92–104.

[14] ——, "Introduction to machine learning with python: A guide for data scientists," 09 2016, pp. 83–91.

[15] S. Krithivasan, S. Shandilya, K. Arya, K. Lala, P. Manavar, S. Patii, and S. Jain, "Learning by competing and competing by learning: experience from the e-yantra robotics competition," in 2014 IEEE Frontiers in Education Conference (FIE) Proceedings. IEEE, 2014, pp. 1–8.

[16] L. K. A. K. Krithivasan S., Shandilya S., "Massive project based learning through a competition impact of and insights from the e-yantra robotics competition (eyrc - 2013)," in 2014 IEEE Sixth International Conference on Technology for Education. IEEE, 2014, pp. 1–8.