RESEARCH TRENDS IN SOFTWARE ENGINEERING FIELD: A LITERATURE REVIEW

Hiba Al Sghaier *1

*1 Yarmouk University, Irbid, Jordan

DOI: https://doi.org/10.29121/ijetmr.v7.i6.2020.694

ABSTRACT

Software engineering is one of computer science branches, it comprises of building and developing software systems and applications. Software engineering is a discipline that has a constant growth in research in aim to identify new technologies and adopt it in different areas; there is a considerable investment on software engineering trends at the current time due to the availability of mobile technologies. With millions of billions of smart devices that are connected to the internet, all industries around the world are rapidly becoming a technology driven industries.

Software engineers are aware of programming languages that are employed to develop software systems, by applying engineering principles to development process; they can design customized software systems for individual or organizational customers. The new trends in software engineering are numerous, Cloud Computing, machine learning, deep learning, big Data, mobile Computing. Nevertheless, there are many more other research trends in software engineering's field that have been intensively explored and implemented in many different industries. In this paper, authors try to summarize the most fields that are integrated with software engineering recently.

1. INTRODUCTION

1.1. SOFTWARE ENGINEERING AND EDUCATION

Education can be accomplished through many methods, like conversations, lectures, spreading of knowledge through social networking etc. In the educational institutions, the knowledge is obtained from teachers, educational materials, negotiation, discussions and communication. Mobile devices became more widespread in our daily lives, and are progressively replacing traditional computer to perform many processes like playing videos, browsing internet, and reading different files [41]

Mobile learning is considered as style presentation learning technique, it depends on using mobiles as an educational platform. using Mobile learning inside schoolrooms enable students to solve problems in associated ways.

A project was developed to provide a gaming environment to self-motivated students [1], in the developed application; each department is presented as a castle with windows, while each window is considered as a course in that department. Each student creates her/his own account, and the account will be initiated with pre-defined points. The student can add one or more questions and assign set points for it, when other students answer the questions,
then points will be assigned to the student’s account. Answering more questions will enable students to get higher level and marked with distinguished badges [1].

Practicing Pair programming in [2] shows a significant impact on the students’ efficiency and gratification, the study prove that students who practice programming in pairs are more compatible than practicing it individually. The developed research in [2] was still required to address other factors that have an effect on Pair programming compatibility; the researchers developed another research in [3] to focus on self-educational mobile and its impact on student’s performance. The developed application provides a game educational environment, the study depends on T-Test to accept or reject some factors that are used to evaluate the developed application, results show that using mobile game applications as an educational technique has a better impact on student performance rather than traditional educational techniques.

A web portal that provides appropriate testing tools (WTLAP) were used in [59]. The portal provided both instructors and students with training materials in the field of testing techniques, and introducing the benefits of utilization such tools in programming courses [59]. Although, educational institutions all over the world have used e learning, the focus of most institutions was on-campus teaching and e learning. In the face of COVID-19, institutions had no enough time to prepare for full transition to e-learning. Study in [4] aims to evaluate the quality of e-learning techniques during the COVID-19 pandemic.

The study [4] proved that institutions in Jordan and Saudi Arabia, were the study was conducted in, still need to improve e-learning quality in many terms like: course sequence, teacher and student preparation and Materials that are used for e-learning. COVID-19 has pushed many institutions to move to e learning, and this allows educators to gain the real importance of eLearning and find solutions to the drawbacks and limitation of the institution’s infrastructure.

2. SOFTWARE ENGINEERING AND MACHINE LEARNING

Software fault prediction is defined as a process of developing modules that help developers to determine the faulty modules or faulty classes at early stages of the development life cycle. Locating faulty modules in any software is considered as an essential step, because it helps to define the modules that need more testing and refactoring. Classification techniques on software fault prediction are widely investigated due to the necessity to decrease testing efforts and increase the software quality [11], [13], [25], [21], [22]. findings of previous researches done on this field has not shown which classifier is superior to the other, a comparative software fault prediction experiments in [10] integrate neural networks and genetic programming [52] performed on 4 datasets from the NASA Metrics Data repository. While in [16] and [15] a symmetric ensemble method is used to predict faulty modules based on the performance of eleven base learners.

In [5], an approach was developed by integrating support vector machine (SVM) with genetics algorithm (GA) and particle swarm algorithm for software fault prediction. The developed approach is applied (12-NASA MDP and 12-Java open-source projects). Results show that integrating the two algorithms with SVM improves the performance of software fault prediction process [5].

In [35], a hybrid machine learning technique and a new feature selection method is used to examine an automated bug triaging system. The developed model used Mozilla Firefox projects for evaluation with respect to precision, F-measure, accuracy and recall metrics. The model was also examined by using boosting, bagging and decorates ensemble methods along with Naive Bayes, Bayes Net, J48, Decision Table and Random Tree base learner classifiers. The results indicate that bagging ensemble methods and decorate improved the classification results by achieved a recall ratio of up to 96%. In [53] experiments were performed with machine learning algorithm (LibSVM classifier in Weka application), in order to classify Quran chapters based on Makki and Madani chapters [53].

3. SOFTWARE TESTING AND TESTING TECHNIQUES

In the software engineering life cycle, one of the crucial activities is to discover and fix errors, and that is done in testing activity, testing activity aims to find bugs and generate test cases accordingly [48].

Software testing is the fundamentally approach that is used evaluate software systems under development. The main purpose of testing is to detect defects before consumers find them out. Therefore, minimizing the cost of the test is a huge challenge [24], [44]. Work in [36] aims to minimize the cost of testing by eliminating the redundant test cases by applying the K-means Clustering algorithm. The results proved that the proposed technique reduced
Research Trends in Software Engineering Field: A Literature Review

the number of test cases from 776 to 240 with keeping the same coverage [36]. Research in [20] focuses on finding a suitable technique to estimate the effectiveness of test cases, Design patterns establish a revolution in software engineering field, as they focus on the importance of reuse and its impact on software product quality and software process [43].

In order to predict the number of test cases, an approach was developed in [12]; it suppose that complex methods comprise more errors, which needs executing more test cases, [14] and [23] examined the ability of complexity and size metrics to determine the most complex and error-prone classes. [18] suggested a concept; it consists of combining different testing patterns into similar test suite to test various aspects through running one single test exactly once.

Pair programming techniques is used for writing developing test cases in order to test software systems. [17], while Information Retrieval (IR) technique is used in [19] as a purpose of test cases’ reduction.

In [49] work, software metrics were used for test case selection, the ability of complexity and size metrics were examined to locate the most complex and error-prone classes, the experiments focused on systems that were written in Java language and tested with the JUnit testing framework. The results indicate that the proposed approach detected most of seeded errors and reduced the needed number of test cases; systems that are written in java language were also used in [51] to evaluate the quantify and efficiency of the developed approach that aim to compare existing mutation-testing strategies [51].

A systematic literature review was conducted in [50] to classify the existing testing techniques and perform kind of comparison in terms of pros and cons. The main result of the developed research is to categorize test suite reduction, which help to provide software testers a guideline for choosing the most appropriate techniques based on the test requirements [50].

4. SOFTWARE ENGINEERING AND DEEP LEARNING ALGORITHMS

Deep Learning is a set of techniques that are used for learning from Neural Networks multi-layers, it is also considered as a subfield of machine learning that uses supervised/ unsupervised strategies. [28]

Previous section talked about using Machine learning in software faults prediction, Deep learning also used in various areas and achieved remarkable performance in areas such as speech recognition, computer vision and natural language processing. Extensive research efforts have been done in the area of Natural Language Processing and one of these researches focused on developing reading comprehension Question Answering systems [54].

In [26], deep learning algorithms were investigated, such as Convolutional Neural Network (CNN) and Multi-layer perceptron's (MLPs); the study addressed the factors that have an impact on the performance of these algorithms. The experimental results show that the impact of modifying parameters had a significate role in prediction performance, which achieved a higher rate in comparison with the Machine-learning algorithm. To validate our assumptions, the result also shows how the estimated factors might decrease or increase the measurement of fault detection rate.

Authors in [27] addressed the susceptibility of employing deep learning algorithms such as (CNN) in the area of predicting students' performance. Different parameters were manipulated to measure the performance in terms of F-measures, Recall and Precision, the study presented promising results in comparison with previous related works results.

A comparative study [29] addressed the effectiveness of deep learning algorithms on Predicting Daily Living Activities for elderly people, (CNN’s) and other traditional machine learning algorithms are employed to perform the experiments. Many factors in the CNN are addressed to obtain the best results in predicting the Daily Living Activities in comparison with the other machine learning results. The dataset was collected from 15 participants performing seven main activities, results prove that (CNN) provides promising result in term of predicting Daily Living activities for the most common datasets.

5. BIG DATA ANALYTICS AND DATA SCIENCE

Big Data became the most important trend in the last few years; Big Data is a technology for (Storing, Processing and analyzing data), many famous organizations and companies like Facebook, Google, eBay, yahoo, twitter and other concern about big data due to its importance in their revenue.
Hiba Al Sghaier

Companies can get the benefits from collecting such data, might be used it to improve their business, products and increase the number of customers by matching their interests and needs through following their feedback. However, big data get along with data analytics; the rapid flow of data means that it should be sorted and processed quickly, Hadoop is a software that have the ability to sort tremendous types of data [30].

Authors in [31] studied the benefits of the utilization of big data analytics in industrial and business procedures can consolidate the agility in the organizations. The transition toward big data analytics can help predictors and decision makers to employ data in taking many actions when striving the intended goals, when organizations adopt big data analytics, they can better predict already unpredictable issues, and improve the processes performance. Big data analytics capabilities could speed up huge amount of organizational operations and process.

Organizations that employ cloud computing and big data analytics use private or public computing networks to handle their work processes, organizations also can share resources within the same organization or share external resources to handle their business at a minimum cost. In [32] researchers conducted an empirical study of big data analytics and cloud computing to highlight the main issues in this field; the study analyzed the previous developed techniques and discussed benefits and challenges of these technologies.

Although the growth of using the data in web-based applications has tremendously increased in the last decades, and these applications created huge benefits to society, yet they still suffer from several security threats, a large number of such applications suffer from wide types of attacks, and result in financial loses [47]. The authors in [52] proposed a method that focuses on improving the effectiveness of the technique of categorizing threats by using (OWASP) that aim to cover most known security attacks. Authors in [57] implied that there is a need to an extensive researches and recommendations for user's awareness to arm them well to the struggle with spam emails and cybercrimes. A vulnerability assessment were conducted on a number of university’s websites in Jordan, and a Passive penetration testing methods were employed. Results showed that a considerable number of the evaluated universities had critical level of vulnerabilities [58].

6. OPEN SOURCE SYSTEM STUDIES (OSS)

Open Source Software is a free software or source code that is available for use and modification. Open Source Software is created by programmers, and a collaborative experience of developers who develop, improve and test these systems. People prefer using open source software systems because it can be found free or with a lower cost. Software maintenance operations may cost 60-77% of the overall cost of software system development [13]. Authors in [33] empirically explored three open source software systems with their eleven versions, versions are developed over time and might have more sources of test suites. In [33] work, mutation testing coverage, code coverage and test suite LOC are measured in order to estimate the correlation between the effectiveness of, code coverage, fault detection, and test suite size. Based on the result authors indicate that "the assumption that test size is always revealing a positive correlation with its effectiveness" could not be generalized.

In [34], the relationship between LOC, nesting levels, and cyclomatic complexity metrics are measured through several software versions, in order to figure out how to calculate the strong relationship between these factors [34]. Sometimes the quality attributes cannot be directly measured. In [37] authors divided attributes into 23 sub-attributes, and then a comparison has been done with the other software quality common models.

7. DESIGN AND MODELING

The changeover from object-oriented (OO) software requirements into design is a fundamental activity in the software engineering processes. Class diagram combines the component's presentation, and it requires much more efforts and focus to be developed. Several developed software researches had made excessive effort in this matter. The approach in [42] based on meta-models, test models and structured models, to reduce the number of test cases in regression testing process. [42], for constructing models of systems, the Unified Modeling Langue is being used as a backbone in software design [45].

Paper in [38] suggested using natural language processing techniques to extract data from use case specifications and translate it into a class diagram. The developed approach identified classes, how they are related to each other, and provided information classes’ internal structure. In [40], authors proposed a model based on testing strategy, which involves four distinct diagrams in the testing process. In order to avoid documentation’s problems in design level [39] applied traceability between design and code, the developed approach support
Research Trends in Software Engineering Field: A Literature Review

traceability between design and code for the Oracle designer tool. In order to evaluate the developed approach, a student information system was used. The results indicate that the developed approach is effective in code-design traceability.

A model-driven approach was used in [46] to map the structural adaptations in autonomic software, metamodel and a workflow were provided to support the approach, a prototype was also developed to imitate a reductive change to an autonomic, service-oriented healthcare application. Big data were used in [56], where a smart system was built to control the lights in Yarmouk university museum. The system consists of thermal sensors that detect the presence of humans and DC LED spotlights. The measurements show the reliability and efficiency of the developed smart system [56].

SOURCES OF FUNDING

None.

CONFLICT OF INTEREST

None.

ACKNOWLEDGMENT

None.

REFERENCES


Mohammad Akour, Izzat Alsmai, MamounAlazab, Malware Detection Challenge Between Accuracy and Performance, 2016 2nd International Conference on Open Source Software Computing (OSSCOM). IEEE

Mohammed Akour, BouchaibFalah, Application Domain Oriented Software Readability Yardsticks, The 7th International Conference on Computer Science & Information Technology (CSIT) Amman-Jordan 2016. IEEE


Mohammed Akour, Hibal Alsghaier, Osama Al Qasem, The Effectiveness of Using Deep Learning Algorithms In Predicting Students Achievements, Indonesian Journal of Electrical Engineering and Computer Science, Volume 19, No 1, 2020


Research Trends in Software Engineering Field: A Literature Review


[34] Mamdouh Alenezia, Mohammad Zaroura, and Mohammed Akour, Exploring Nesting Levels in Evolving Open Source Software Systems: Experimental Study, Accepted (1/2019) for publication in Recent Patents on Computer Science.


[37] Sultan T. Alanazi, Mohammed Akour, Mohammad Anbar, Enterprise Resource Planning Quality Model ERPQM, First International Conference of Intelligent Computing and Engineering (ICOICE 2019), Yemen on December 15-16, 2019 IEEE.


International Journal of Engineering Technologies and Management Research
Hiba Al Sghaier


