



Web Application to Generate and Classify Random Walks of The Particles Using Machine Learning

Afrah Mahmoud Mousa*

Department of Software Engineering, Bethlehem University, Bethlehem, Palestine

***Corresponding Author:** Afrah Mahmoud Mousa, Department of Software Engineering, Bethlehem University, Bethlehem, Palestine.

Received: January 25, 2023

Published: May 04, 2023

© All rights are reserved by **Afrah Mahmoud Mousa.**

Abstract

The subject of random walks of particles is taught in universities and schools in a theoretical manner for several reasons, the most important of which is the lack of sufficient programs to teach such subjects, which makes the process of teaching such subjects a difficult task for teachers, due to the lack of a convenient and appropriate way to reach students with the correct idea of the material. Which causes a drop in the educational attainment rates for students in the science courses that contain the subject of the random walks of the particles. In this project, artificial intelligence and machine learning techniques have been used to solve this problem, which uses to help in teaching and studying the subject of particle random walks.

Keywords: Random Walks; Machine Learning; Model; Application

Introduction

This project, which is entitled “Classify Random Walks Using Machine Learning”, combines two important fields, the first field is science such as physics, mathematics, biology, and the second field is computer science, so this project is a web application that allows users, by machine learning, to classify a set of random walks of the particles, and plot them using 2D plots.

Three machine learning supervised algorithms were used in this project: K-nearest neighbors algorithm, Gaussian naive Bayes, and decision tree. The users have the ability to choose between these three models to do the classification. Several programming languages were used for the project, including python3, HTML/CSS, JavaScript, and SQL for the project database.

This project has been done for three customers after doing initial interviews with them: Bethlehem University, Jülich Research Institute, and some high schools in Germany. Where some of them were interviewed face to face, and others through the Zoom application, in the interviews they were asked about the problems they suffer from in the subjects of random walks and machine learning.

A set of questionnaires has also been published on the university's and the institute's official websites to take a larger sample of users.

Moreover, the requirements that were not clear at the beginning were clarified by involving the users in the development process and getting their reviews and feedback after completing each step.

Teachers, students, also researchers from these organizations can use the application for their work.

Background

There is a famous sentence that says: “Tell me and forget. Teach me and I may remember. Involve me and I learn” [1], this sentence was the biggest motivation for doing this project. From being a student, I know very well that education without practice and experiments does not benefit the student at all, and of course, it does not benefit the teacher as long as the students cannot understand the content, so from this, I started the idea of the project, which focuses on the principle Learning by doing.

Artificial intelligence has recently spread in the world and has dominated the electronic products field because it contributes to making human life easier, and for this reason, machine learning has been chosen to use for this project, in addition to being easy to use to accomplish such projects.

Choosing the subject random walks of particles is due to the fact that this subject is taught in science departments in universities, scientific institutes, and some high schools in a theoretical not practical way, and this is what made the task of teaching it difficult for teachers also for students because there is no practical thing explaining the material to facilitate its understanding.

This traditional method of teaching is not suitable for teaching such subjects. For example, the student needs something tangible and practical to notice the difference in the change of random walks when changing the angle, for example, or the length of the walk, in order to conclude the relationship between them. Words and pictures do not clarify this purpose at all, and therefore the idea of the relationship between the angles, walk length, and the random walks is not properly understood by the student, so the student gets low grades.

Also, the teacher needs a practical and tangible application to clarify the difference and the relationship between them, not pictures and words that everyone can memorize and repeat.

The same applies to the subject of machine learning, as computer subjects in general, can only be taught using practical applications, as is well known.

And after making questionnaires And the interviews figures that students and teachers need such an app to help them.

The application is not only limited to science students and teachers but also computer science students and teachers can use it for machine learning subjects because the app contains several machine learning models that allow them to make comparisons between them in the results, actually, the application itself is based on machine learning.

Project functionality

This application performs three main tasks, which can lead to giving secondary tasks as part of them, these three tasks are.

Generate new random walks

The application can generate new random walks by writing the attributes for each random walk by the users in the correct place inside the application. These attributes are previously defined by the admin, but their values are entered by the users, each user can enter his own values.

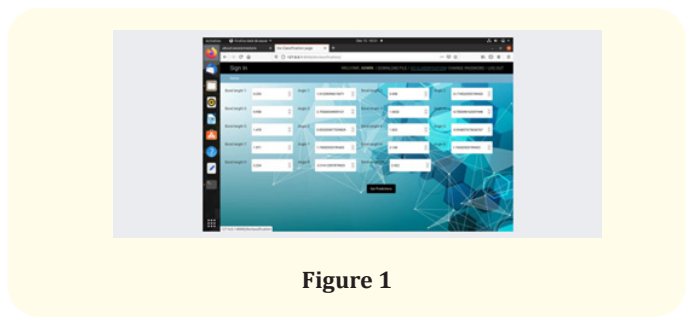


Figure 1

Plot the random walks

The application will present to the users the plot of the random walk trajectories that they generated before by them, so they can see how the particles moved according to the information they give before.



Figure 2

Classify the random walks

The app will classify the random walks that the users generated to the correct class according to the user inputs.

As for the secondary tasks that the application does, they are:

Allow selection of the classification model

The users can choose between the three machine learning models to do their classification.



Figure 3



Figure 4

Save the plot figure

The users can save the random walks plots on their devices.

Download the classification result of the default dataset

The application contains the default dataset that the admin provides, the users can classify these random walks which are inside the dataset file, and download the results file in their devices.

Approach and methods

In this project, Agile methodology was adopted for the software development, because it is flexible, fast, and keep the project updated with the environment update.

Agile methodology helped to deliver the project functionalities in the time, as it is fast which means it takes less time for the development rather than the other software development methodologies, also it enforces us to focus on the project quality not the time, to give the customers qualified services.

Agile Methodology depends on continuous communication with customers, and this is what has been done in this project, there were weekly meetings were held with customers in order to record their reactions and reviews about the application, then these reviews and reactions were used to improve the application quality and develop what the customers might like in the application. All this led to customer satisfaction with the application.

These are the Agile methodology process steps that were followed to gain the Agile development values:

- **Planning:** In this step, the requirements of the customers were taken and classified, in addition to the identification of the specific customers for the project, and the technologies that will be used for doing the project.
- **Designing:** Here the software development diagrams were made that explain and clarify how the program works, such as workflow diagram.
- **Developing:** This step is in which the whole program was developed, but before the development, all the software languages and tools were selected to do the project development, including the operating system, the IDEs, and more.
- **Testing:** Here, what has been accomplished from the project was presented to the customers, their reactions, reviews, and suggestions were recorded, for using it in developing the project next time.
- **Deploying:** This step is the last stage in which the project has been completed and released to the customers.

Upon completion of all steps, we went back to the initial step and start again, as the project was divided into smaller parts to facilitate work on them and to conduct work faster, and this is what the Agile method depends on.

Also, Agile Testing was used for the project to ensure that the project is running correctly and doing all the functionalities right without errors or bugs, as new tests were performed whenever new features were placed in the application, and not after the completion of the development phase.

Acceptance test-driven development (ATDD) has been used for testing the project by first writing the user acceptance test which had been collected from the users and then writes the code until it passes the test.

For the machine learning models, cross-Validation has been used to verify the validity of the machine learning models that were used in this project and to ensure that it works without errors and with high accuracy with the number of folds = 5.

Experiments

A sample of random walks data was taken from the Jülich Research Institute, which was worked on during this project.

Several models of machine learning models have been used, as they were trained by the training data and then tested by the test data. The data was divided into the test data and the training data from the original dataset file that contains 1000 records, where 70% of the data was calculated in the file to training data and 30%, which constitutes the remainder of the file, was calculated for the test data. These percentages were chosen in a logical, appropriate and actual manner.

Five types of machine learning models were worked on, but in the end, three of them were chosen based on several measures, which they are: AUC, recall, precision, and F1. These measures were as follows:

| Model | AUC | Recall | Precision | F1 |
|---------------------|-------|--------|-----------|-------|
| KNN | 0.983 | 0.893 | 0.898 | 0.894 |
| SVM | 0.501 | 0.317 | 0.361 | 0.273 |
| Decision tree | 0.997 | 0.989 | 0.989 | 0.989 |
| Naïve Bayes | 0.903 | 0.784 | 0.785 | 0.783 |
| Logistic Regression | 0.665 | 0.578 | 0.597 | 0.502 |

Table 1: Machine learning models measures.

From these measurements, the KNN, Decision tree, and Naive Bayes have been chosen to do the classification in the application [2,3].

Conclusion and Recommendations

This project will contribute significantly to improving the educational process as it combines theoretical and practical education, this problem is one of the problems that we are currently facing in schools and universities, especially in teaching a subject such as the random walks of the particles that the project focused on, as it requires seeing the walks and recording notes repeatedly, and this project came to solve this problem, and we will see clear progress in the academic achievement of students after using the application.

Teachers and students in the field of machine learning can also use it because it makes it easier for them to understand how the classification works, and they can do the classification comparison as the application has three machine learning models.

Not only teachers and students who can use the application, but researchers in the field of particle random walks can use it too, as it helps them in their work by writing the attributes of each random walk and knowing the class to which it belongs.

My suggestion to make the project achieve success is to include the application in schools and universities on their devices, and to use this application as an essential part of teaching such subjects.

It is also preferable to give teachers and students a half hour or more class explaining to them how to use the program before starting to use it, so that it is easier for them to deal with it and understand how it works.

Bibliography

1. "Chinese Proverb".
2. IEEE Conference Template Example Overleaf, Online LaTeX Editor.
3. What is Agile Methodology? Tools, Best Practices and More, Website title: Stackify.