|  |  |
| --- | --- |
|  | |
| **Department of Engineering**  **Higher Diploma in Computer and Electronic Engineering**  **Artificial Intelligence and Big Data (EEE4463)**  **PBL Project Brief:**  **Data analytics: Insight extraction from projectile motion** | |
| **AY2021/2022** | Engineering Discipline |

1. **PBL Project Title**

Data analytics: Insight extraction from projectile motion

**Required component list**:

ParabolaWithNoiseOutlier.csv

LinearRegression\_ParabolicTest01.py

1. **Project Overview**

You are a data analytics engineer testing a cannon prototype. The prototype is built for testing the performance if the cannon meets the design requirement. You are going to analyze the projectile motion of a cannonball fired by the cannon prototype. The positions of the cannonball at different time instants are measured by a RADAR. However, the data is corrupted with noise and outliers. It is necessary to condition the data before building a model describing the projectile motion of the cannonball. Finally, we try to extract as much information as possible from the model built.

We have a practical scenario that a cannon is placed above a cliff firing a cannonball. We have a RADAR capable of measuring the positions of the cannonball at different time t. When the cannon fires a cannonball, the RADAR keeps measuring the positions of the cannonball at different time and saves them into a file “ParabolaWithNoiseOutlier.csv”. The first and second column of the csv file represents the x and y coordinates of the position of the cannonball respectively. The unit of the x and y coordinates are meter. The origin of the coordinate system is the position of the cannon as shown in the following figure.

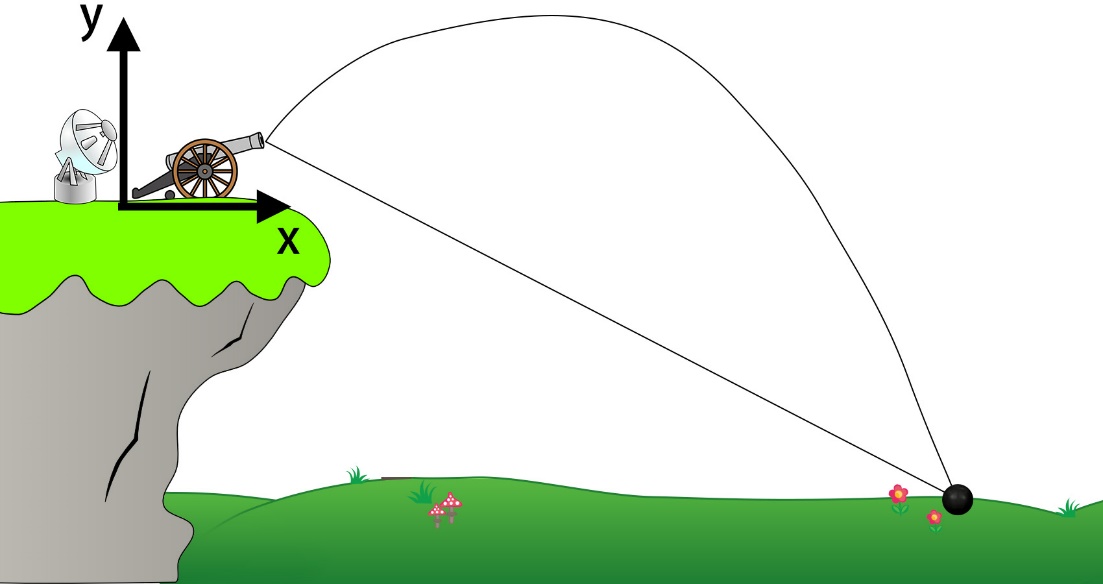


Figure 1 testing scenario for the cannon prototype

We have to analyze the data and design the procedures to process the data using appropriate tools. Next, we build a model for the data, analyze the model and extract as much information from the data. Finally, write a report to illustrate (i)your designed procedures and tools you used to process the data, (ii)the method you used to build the model, and (iii)the information you extracted from the data and model. Provide all your Python programs used at the end of the report as appendix.

A Python program with filename “LinearRegression\_ParabolicTest01.py” is provided demonstrating the way to build a 2D model using linear regression method.

The following is the guideline to design your program, analyze the data and write your report.

1. Use the program “LinearRegression\_ParabolicTest01.py” to process the data in the file “ParabolaWithNoiseOutlier.csv”.
2. Briefly explain what the program does.
3. From the results and figures, comment on the results, the condition of the data and suitability in using the linear regression method to build model.
4. Design a new program using appropriate tools to process the data.
5. Briefly explain what the new program does.
6. From the results and figures, comment on the results and suitability in using the tools to build model.
7. Challenging part: Make use of the knowledge of moving object with constant acceleration, work out an equation to describe the relationship between the x and y coordinates.
8. Challenging part: Compare your equation with the estimated model and extract as much information from the estimated model.
9. Challenging part: Explain the difference between linear regression method and RANSAC method
10. **Driving Question**

As a data analytics engineer, how do we condition the noise corrupted data? How do we build a model to model the projectile motion? What insight can be extracted from the model built? How do we use this new insight?

1. **Target Users**

* Data analytics engineer
* Data scientist
* Cannon operator
* Cannon designer

1. **Project Structure**

Each team should involve 2 students participating in this PBL project. Students are recommended to work with the following framework to analyze the provided cannonball positions data. Students need to determine which method to be used to condition the data (e.g. to remove outliers) and fit a regression model on the conditioned data. Subsequently, students need to extract information from the model built. Finally, give some suggestions in using the new insight extracted.

* Analyze the provided position data of the cannonball
* Review the topic of linear regression in this module EEE4463 Artificial Intelligence and Big Data
* Search the internet, ebooks, textbooks, or any other means for key knowledge including
  + Motion with constant acceleration: the equations describing object motion under constant acceleration
  + Linear regression: a method to regression analysis which fits a regression model on data
  + RANSAC regression: a better approach to regression analysis which fits a regression model on a subset of data that the algorithm judges as inliers while removing outliers.
* Note that we can assume there is no air resistance applied to the cannonball
* Determine which artificial intelligence method to be used to remove the outliers
* Extract insight based on the model built
* Review, voice out any questions and give feedback to and get feedback from teachers and peers in every step.

1. **Inter-disciplinary Learning and Modules Support the Project**

This Project comprises intended learning outcomes of the following modules:

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Credit** | **Module Code** | **Name of Module** | **Modules’ Intended Learning Outcomes (MILOs) which contribute to the Project** | | | | | **Deliverable(s) from each module for the Project** |
| MILO 1 | MILO 2 | MILO 3 | MILO 4 | MILO 5 |
| 14 | EEE4463 | Artificial Intelligence and Big Data | -- | -- |  |  |  | Students need to determine which A.I. method to be used to condition noise corrupted data. Based on the model we built, extract insight. |
| 6 | LAN3100 | E & C: Workplace Interaction | -- |  | -- | -- | -- | Students need to make suggestions by giving opinions with justification appropriately. |

1. **Roles and Responsibilities**

Students as active learners manage this student-led project and take greater responsibilities to learn the key knowledge and success skills corresponding to the driving questions.

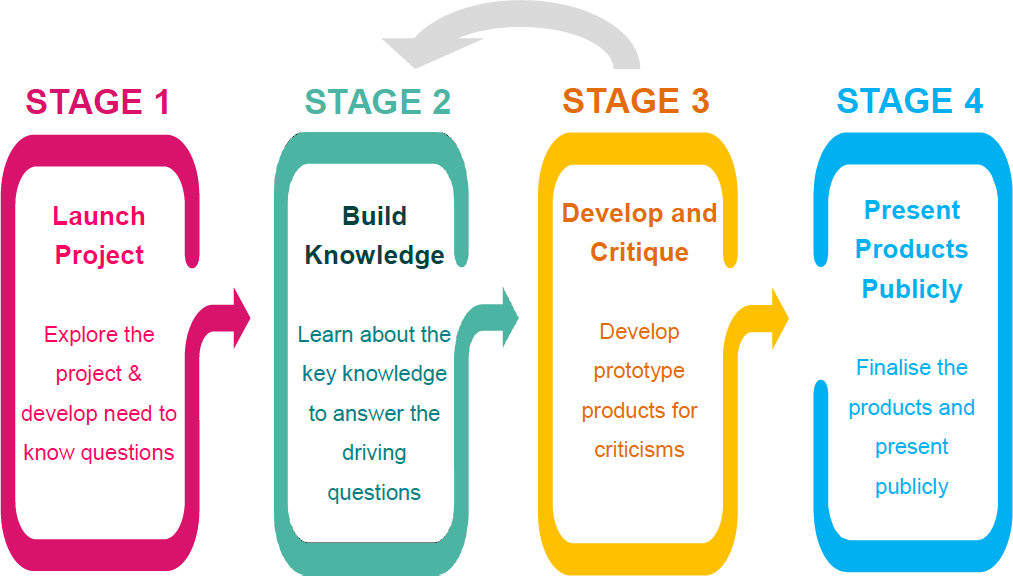
In doing this project, the role of students is the engineer to analyze the acoustic performance of the air-conditioning system and suggest solutions when problems are found.

Students’ responsibilities include but not limited to:

* Find resources & apply information
* Use standards/code of practice
* Make decisions on how you work
* Concern environment, safety, and sustainability
* Demonstrate your engineering competence and professional competence
* Present your deliverable in a logical and easy understanding format and prepare to present it publicly
* Apply 21st Century skills, particularly critical thinking, communication, literacy skills, flexibility, initiative, productivity, and social skills for this project
* Keep your classroom behavior so that the classroom norms can maximize the productivity of every student
* Proactively seek for and apply feedback from self-assessment, peer and teacher, etc.. Instructional supports are available but these supports are removed when no longer needed. Note that teachers will not “front-load” too much information at the start of the project, but waits until it is needed or requested by students
* Keep enthusiasm and sense of ownership of the project, commit to the project calendar and make improvements on the products of the project
* Communicate thoroughly with teachers and stakeholders using digital communication method

1. **Project Path and Milestones**

The Project is for the module Artificial Intelligence and Big Data in semester 3 (Full-time study). The Project development is divided into four stages as shown below:



Project Path for PBL (PBLWorks (n.d.))

|  | **Stage No.** | **Learning activities** | **Submission** |
| --- | --- | --- | --- |
| Week 40  OR  At the time start the topic | 1:  Launch Project | * Initiate the inquiry process through participating in a project briefing by the end user * Preliminary study the problem to be solved in a group of 2 students. * Figure out the solution to the problem. Plan the sources; resources include but not limited to the key knowledge, available time, available knowledge and skill sources such as websites, books, ebooks in library. | -- |
| Week  41-42 | 2:  Build Knowledge | * Study the methods in conditioning data and determine which methods is the best to remove outliers. * Built a model based on the conditioned data * Study the motion with constant acceleration to figure out an equation to describe the model. | Verbal submission by each group to show the progress each week |
| Week  43-44 | 3:  Develop and Critique | * Carry out the analysis for the problem * Peer review the analysis, * Give, receive and apply feedback * Continue to carry out the analysis and draw conclusions, think about the implication, and provide solution to solve the problem | Verbal submission by each group to show the progress each week |
| Week  45 | 4:  Present Products[[1]](#footnote-1) Publicly | * Present the work in written form and prepare to present it publicly. * Conduct self-evaluation of the work and performance, and reflection on what have learned in the project | * Presentation of the extracted insight and the method used * Report submission to Moodle |

1. **Professional Ethics**

**9.1 Non-disclosure Agreement/ Confidentiality**

Students should observe absolute confidentiality in all matters concerning a proprietary technology, strategy, organization and business practices, and/or any other matter which is defined as confidential unless permitted to disclose such information by the owner, particularly the information and technical data in the product catalogue or service manual.

**9.2 Intellectual Property and Plagiarism**

Students must note clause 3.10 of the Student Handbook and shall not violate the VTC’s Intellectual Property Policy for Students.

Plagiarism is strictly prohibited. It is an academic offence, but also being recognised as a violation of copyright law when the act of plagiarism involves others’ Intellectual Property. The Student Handbook clause 3.11 “Guiding notes on using or copying of copyright works” and clause 7.10 “Cheating and irregularities relating to electronic / communication devices during assessment give guidance for students.

**9.3 Digital Ethics**

References: GovHK, March 2021. Cyber Ethics. [online] Available at: <https://www.gov.hk/en/residents/communication/infosec/cyberethics.htm> [Accessed 18-04-2021 16:00]

9.3.1 Data Protection

The Personal data of you and the parties involved in this project must be

* Processed lawfully, fairly, and in a transparent manner
* Collected for specified, explicit, and legitimate purposes
* Accurate, adequate, relevant, and limited to what is necessary to fulfill the purposes for which it is processed
* Kept for no longer than is necessary for the purposes for which it is processed
* Processed in a way that ensures it is appropriately secure

9.3.2 Cyber Security

* All computer equipment shall not be left unattended
* Backup shall be performed regularly to ensure that system and data can be recovered from the disaster
* Email services provided by the Council shall only be used to support the operation of the Council. The services shall not be used to perform personal activities or other illegal activities such as sending spam email
* Students shall take note of the “Terms of Use”, “Cookie Policy” and “Privacy Policy”.

1. **Safety Precaution**

**10.1 Psychological and Emotional:**

* If psychological and emotional problems are encountered when handling the assignment as it will count as a part of your module mark. You are strongly advised to look for support from colleagues in SDO, Department, and any other persons you think suitable.
* Other resources: Department of Health The Government of the HKSAR, *Emotional Health Tips,* [online] Available at: <https://www.studenthealth.gov.hk/english/emotional\_health\_tips/txt\_emt\_mental\_health\_resources.html> [Accessed 18-04-2021 16:00]

**10.2 Environmental:**

* Minimize the use of any materials when doing this project and assignment

1. **Assessment Marks Distribution**

The assessment components and respective weightings as below:

|  |  |
| --- | --- |
| **Assessment Components** | **Weighting** |
| Demonstration and presentation of the project  The following contents are suggested.   1. Plot the data stored in the provided csv file 2. Briefly describe the data and problem based on your analysis 3. Describe and explain your proposed method to condition the data 4. Describe and explain the steps to build a model on the conditioned data 5. Show the code, diagrams and the result after processing the data using your method 6. Describe and explain your insight | 50% |
| Final report  The following contents are suggested.  1. All the contents in demonstration and presentation as stated above  2. Show the steps in obtaining an equation describing the projectile motion  3. Compare your equation with your model built  4. Describe your insight based on your analysis  5. Suggest an application based on your new insight  6. Explain the difference between linear regression method and RANSAC method | 50% |

The below table summarizes a breakdown of individual and group assessment:

|  |  |  |  |
| --- | --- | --- | --- |
| **Assessment Components** | **Weighting** | **Group Assessment** | **Individual Assessment** |
| Demonstration and presentation of project | 50% | 70% | 30% |
| Final report | 50% | 70% | 30% |
| **Total:** | **100%** |  | **100%** |

1. **Assessment Rubrics / Marking Scheme**

Refer to the assessment marking scheme.

|  |  |  |
| --- | --- | --- |
| Grade | Mark Range | Description |
| A | 75-100 | *Knowledge Application*   * Produce an equation describing the model. All of them are correct. * Information extracted from estimated model are all correct * Produce and design a new program to process data which is outstanding * Provide outstanding explanation to the new program * Provide outstanding explanation to the provided program * Provide outstanding analysis of the results and figures generated by the provided program * Provide outstanding comment to the suitability of using linear regression to the provided data   *Report Writing Skills*   * Complicated concepts and ideas are presented professionally and logically with outstanding arguments * Various views and analysis are critically evaluated and synthesized * Creative and innovative recommendations are proposed in the report * The format and structure of the report is presented in a neat, professional and consistent style |
| A- | 70-75 | *Knowledge Application*   * Produce an equation describing the model. Most of them are correct. * Information extracted from estimated model generally correct * Produce and design a new program to process data which is reasonable * Provide reasonable explanation to the new program * Provide some explanation to the provided program * Provide some analysis of the results and figures generated by the provided program * Provide some comment to the suitability of using linear regression to the provided data   *Report Writing Skills*   * Complicated concepts and ideas are presented professionally and logically with convincing arguments * Various views and analysis are critically evaluated and synthesized * Creative and innovative recommendations are proposed in the report * The format and structure of the report is presented in a neat, professional and consistent style |
| B+  B  B- | 60-69 | *Knowledge Application*   * Produce an equation describing the model. Some part may be incorrect. * Information extracted from estimated model may have some error * Produce and design a new program to process data which is reasonable * Provide reasonable explanation to the new program * Provide some explanation to the provided program * Provide some analysis of the results and figures generated by the provided program * Provide some comment to the suitability of using linear regression to the provided data   *Report Writing Skills*   * Rich information/ ideas are communicated with professional purpose and message * Good investigations and analysis are recorded and presented in creative way * Constructive conclusion and recommendations are drawn in the report * The report is wells-structured and organized with sections and sub-sections |
| C+  C  C- | 50-59 | *Knowledge Application*   * Fail to complete Challenging parts * Fail to extract information from estimated model * Produce and design a new program to process data which may have some problems * Provide some explanation to the new program * Provide some explanation to the provided program * Provide some analysis of the results and figures generated by the provided program * Provide some comment to the suitability of using linear regression to the provided data   *Report Writing Skills*   * Reasonable information/ ideas are communicated with well purpose and message * Relevant investigations and analysis are recorded and presented in attractive way * Clear conclusion and recommendations are drawn in the report * The format of the report is well-organised and professionally |
| D+  D  D- | 40-49 | *Knowledge Application*   * Fail to complete Challenging parts * Fail to extract information from estimated model * Fail to design a new program to process data * Superficially explain the provided program * Provide little analysis of the results and figures generated by the provided program * Fail to comment the suitability of using linear regression to provided data   *Report Writing Skills*   * Relevant information/ ideas are communicated with a clear purpose and message * Limited investigation and analysis is recorded and presented in clear way * Only simple conclusion is drawn in the report * The format of the report is organized and complies with the requirements and specifications |
| F | 0-39 | *Knowledge Application*   * Fail to complete Challenging parts * Fail to extract information from estimated model * Fail to design a new program to process data * Fail to explain the provided program * Fail to analyze the results and figures generated by the provided program * Fail to comment the suitability of using linear regression to provided data   *Report Writing Skills*   * Irrelevant information/ ideas are communicated with an unclear purpose and message * Lack of investigation * No conclusion is drawn in the project * The format of the report is messy and without organization |

1. [↑](#footnote-ref-1)