Career Readiness

By Felix Chen and Aarav Rathore

21549A Gearsquad

Queen Elizabeth's School, Barnet, London, United Kingdom

Data scientist

INTRODUCTION

Our team has chosen the career of data scientist as it is an exciting and rewarding career path for individuals with a passion for data, problem-solving, and strong communication skills. Data scientists are at the forefront of driving innovation and making data-driven decisions that impact businesses and society through data analysis, machine learning, and programming knowledge. They are responsible for collecting, organising, analysing, and visualising data to identify patterns, trends, and anomalies, which is crucial in helping businesses make informed decisions, improve their products and services, and gain a competitive edge.

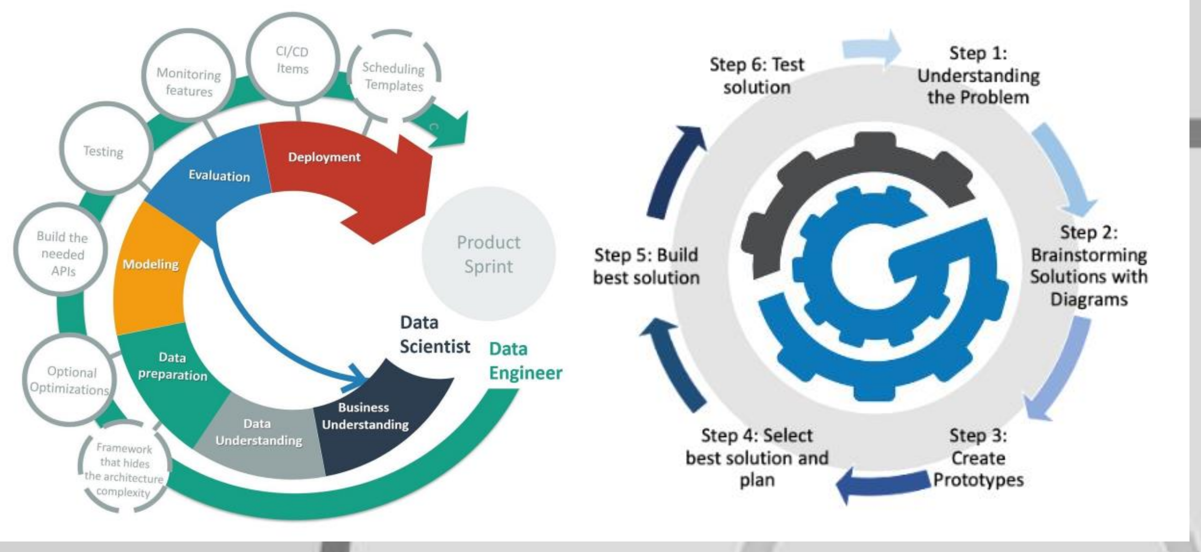


Figure 1: (left) The engineering design process used by data scientists (right) 21549A Gearsquad's engineering design process

INTRODUCTION

Data scientists follow a specific engineering design process that enables them to work effectively alongside other scientists and data engineers, who build and improve systems that help convert raw data into usable information for data scientists.

We have specifically researched the Google Cloud analysis of COVID-19 public datasets, created through Big Query and Data Studio. It provided insights into the pandemic and helped visualise data easily for the general public's ease of understanding. This project involved collecting and organising public datasets from various sources which contained information on confirmed cases, deaths, tests, hospitalizations, mobility, and economic activity, which were available for free on BigQuery, a serverless data warehouse that allows users to run fast and scalable queries on large volumes of data. Other tools such as Cloud Datalab, Cloud Dataproc, Cloud Dataflow and Machine Learning were also used in analytical processes, enabling the exploration, preprocessing, and transformation of data used to join the COVID-19 data with Google Community Mobility Reports to examine the relationship between social distancing measures and the spread of the virus and forecast the future trends of the pandemic.

STEP 1: UNDERSTANDING THE PROBLEM

First, there is an understanding of the problem or task. This is usually done during a meeting and through a brief that tells the data scientist everything they must complete and all of the requirements that their model must meet by the end of the design sprint. Our design cycle begins similarly, with the decision of our goals and what we want to achieve by the end of our design cycle.

State the goal

Begin by stating your goal in non-ML terms. The goal is the answer to the question, "What am I trying to accomplish?"

The following table clearly states goals for hypothetical apps:

Application	Goal
Weather app	Calculate precipitation in six-hour increments for a geographic region.
Fashion app	Generate a variety of shirt designs.
Video app	Recommend useful videos.
Mail app	Detect spam.
Financial app	Summarize financial information from multiple news sources.
Map app	Calculate travel time.
Banking app	Identify fraudulent transactions.
Dining app	Identify cuisine by a restaurant's menu.
Ecommerce app	Reply to reviews with helpful answers.

Figure 2: The goals stated on Google Cloud for different applications

STEP 1: UNDERSTANDING THE PROBLEM

Specifications for our robots

However, we need to first define what our robots this season will need to do to achieve maximum points and what we need to work towards.

Our Robot **Must** be able to	Our Robot **should** be able to	Our Robot **could** be able to
Collect green blocks from the ground. Green blocks are the smallest and easiest blocks to intake. They are also the most populous on the field	**Sort of the green and purple blocks apart.** We can achieve bonus points if the all the blocks in the goals are the same colour. We need to be able to sort out the smaller 2 blocks sizes so they they can be put into separate goals.	**Clear the Supply Zone.** Clearing the supply zone is probably one of the most challenging parts of this game. Since the blocks are of different sizes, they are randomly placed and the fact that purple blocks do not fit under the supply zone bar this makes it one of the most hardest parts of the game.
Collect purple blocks from the ground. These are the next-smallest blocks. There are 16 of them on the field.	**Be able to pick up and put red blocks into Goals.** The red blocks are very large. This is good and bad. Firstly, it would be harder to pick up and store these large blocks in our robot, however we only need 2 of these blocks to fill the goal to fill level 3, earning us loads of bonus points.	**Fully Park in the Supply Zone at the end of a match.** Fully parking the robot earns us 10-20 bonus points at the end of the match which would be very useful.
Dispense the blocks into goals. Nearly all of the points which can be gained in Full Volume are by putting the blocks in the goals.		
Knock over red blocks. Red blocks are extremely easy to knock over and for each red block which is knocked over, we earn 5 points!		
Partially park. In partial parking we just have to have the tiniest part of our robot on the supply zone before the end of the match.		

Project Overall Robot Specifications

Name Aviral Goel / Dev Kamesh Date 05/07/2023 Page 12

Specification for Drivebase

To allow our robot to move around we need a drivebase. This season, the drivebase is more important than ever, because it needs to be fast, strong and maneuverable as always, but it also needs to be able to climb over the supply zone bar for the Endgame.

Our drivebase **Must** be able to	Our drivebase **should** be able to	Our drivebase **could** be able to
Drive at a relatively fast speed. Each match is only 1 minute, so every second gained from travelling around the field faster is essential.	**Compact** A compact drivebase allows more space to fit mechanisms which need to interact with the field such as an intake and also to add support on the outside of the drivebase without breaking size	**Fully park.** This may require dedicated modifications to the drivebase to allow it to quickly drive over the supply zone bar. However, if both robots do this at the end of a match, a bonus of 30 points total can the obtained.
Support a lot of weight and be strong. The drivebase must be able to support all of the mechanisms attached onto it, even while moving at high speeds during a match and experiencing sudden shocks or changes in velocity. It should also not bend in the middle or flex under the weight of the robot.		**Move in all directions** An Omni-directional robot would help us move around a field faster, without having to turn, however it would come at the cost of a motor.
Maneuver quickly without compromising stability. This is crucial as it prevents accidental tipping over when suddenly changing velocity, and saves time that may be lost from resetting the robot during a match, as well as any points lost due to any blocks that fall out during a maneuver.		
Partially park This is a quick and easy way to earn points at the end of the match, requiring only a small part of the robot to be in the supply zone, and not needing the robot to climb over the bar. This makes it useful if time is running out.		

Project Drivebase Specifications

Name Felix Chen Date 05/07/2023 Page 13

Specifications for Intake

To be able to collect and store blocks in our robot, we first need an intake mechanism. This intake mechanism needs to be able to take blocks from the floor and put them in our robot.

Our drivebase **Must** be able to	Our drivebase **should** be able to	Our drivebase **could** be able to
Collect green blocks from the ground. Most blocks are green, and they are the easiest to collect. Collecting only green blocks will also mean a uniform bonus is guaranteed.	**Collect several blocks simultaneously.** This greatly speeds up the intake process, allowing more time to be spent on other parts of the match, like dispensing and clearing the supply zone.	**Double as an outtake system.** This would save motors, which could be used on other mechanisms, for example implementing H-drive, which needs one more motor than a regular tank drive drivebase.
Collect purple blocks from the ground. There are less purple blocks, but these are larger and worth more points. In addition, they are still quite easy to collect, and by collecting only purple blocks a uniform bonus may also be obtained. It is also easier to gain height bonuses with purple blocks.	**Collect two or more colours of blocks simultaneously.** This further speeds up the intake process, as it means the robot does not need to spend time collecting one type of block, then the other, and can instead just drive over all types of blocks to intake them at the same time.	**Collect all blocks from the supply zone while only partially entering the zone.** By being able to clear the supply zone with only partial parking this greatly reduces the time needed to clear the zone. This is because there is no need to climb into and back out of the supply zone to clear it.
		Collect red blocks from the ground. These are the largest blocks, and are worth the most points. They are the hardest to intake and store, but can easily obtain height and uniform bonuses.

Project Intake Specifications

Name Felix Chen Date 10/07/2023 Page 14

Figure 3: The specifications stated in our Notebook to be achieved by the end of the design cycle for each mechanism of our robot

STEP 2: UNDERSTANDING THE DATA

After this, there is an understanding of the provided data or data sources. This is usually provided in the brief and further discussed before anything else is done. Before we start building anything, our team also goes through the same process, looking at the available resources and time, which allows for better planning and time management during the design cycle and ensures that there is no running out of parts.

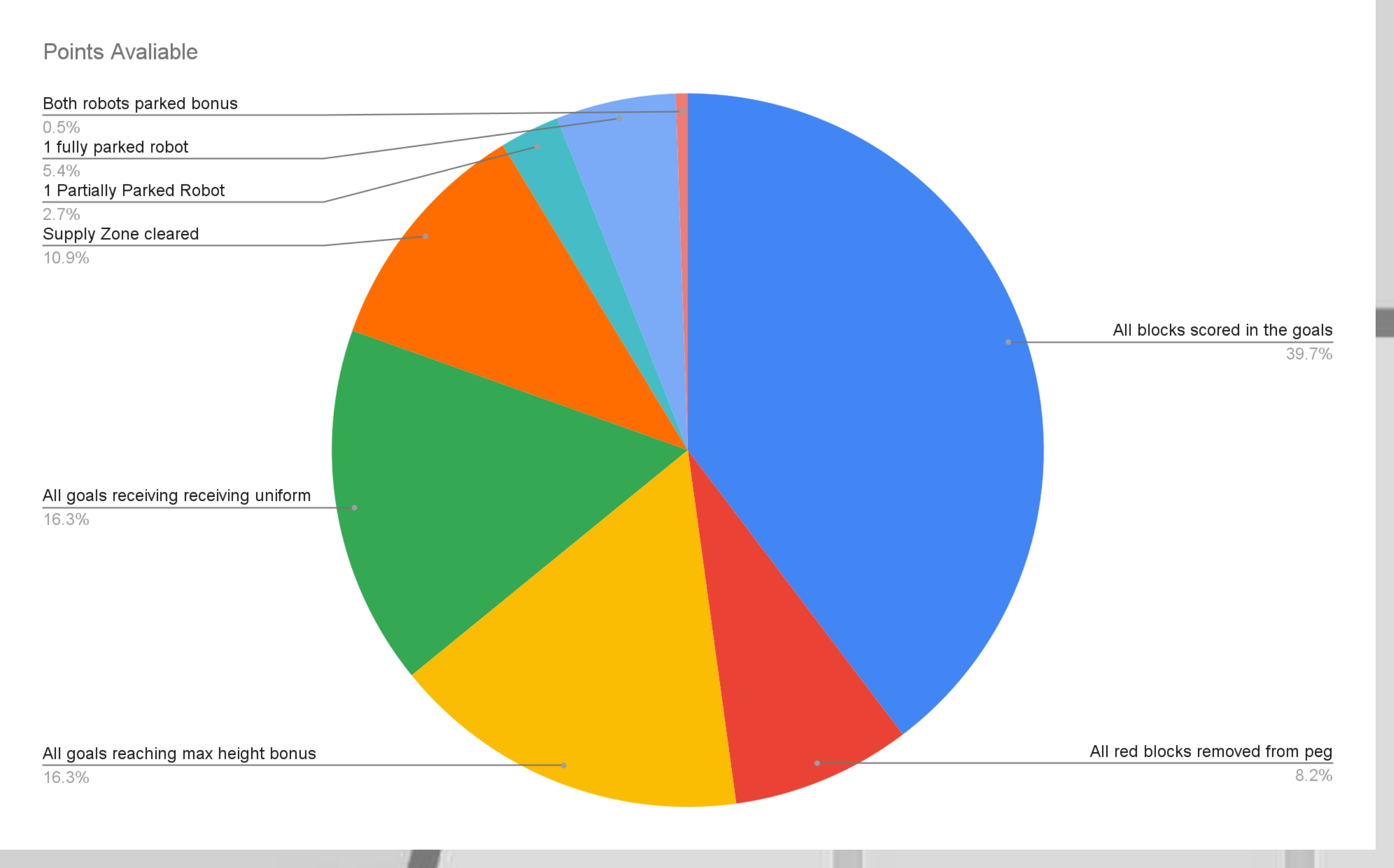


Figure 4: A pie chart indicating the percentage of total points earnt for each task completed in Full Volume

STEP 3: PREPARATION OF THE DATA

Following this is the preparation of this data, which is done through data manipulation and cleaning, which requires tools such as SQL, Python and R. These can automate the task of identifying and collecting data from various sources, including internal databases, external APIs, and online surveys. This prepared and organised data is then used to create models, which is usually done through machine learning and statistical analysis, a process that allows a computer to recognise patterns or specific objects progressively better over time, allowing even the most complex datasets to be analysed in a relatively short space of time.

In this stage, we create and select the best prototypes for each part and mechanism of our robot to be combined later on.

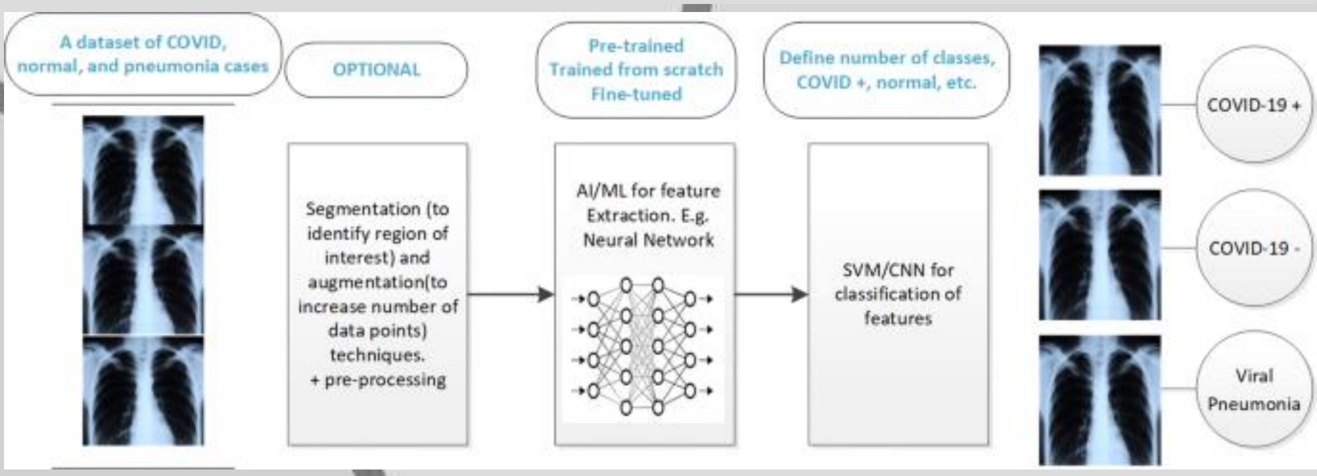
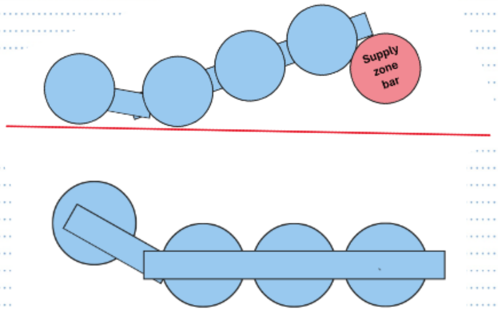


Figure 5: The machine learning model used by Google to create their Google Cloud Covid-19 Dataset to differentiate between Covid-19 and other viral diseases

In our previous design cycle, we only tried basic drivebase configurations. We can now start thinking about making our drivebase more advanced.
We would move the front wheel all the way to the front. Using this wheel, we could climb over the low supply zone bar.
During this, the robot would probably be susceptible to fall over on its back. To avoid this, we thought of putting an extra wheel on the back of a standard drivebase. This wheel would be raised above the ground. When the robot leans back, it would rest on this wheel. The wheel would be connected to the rest of the drivebase, so would be able to also drive the robot forward.
We also thought about adding a 3rd wheel in the middle of the drivebase on each side. This would allow more grip while going over the bar.
Furthermore this should not make any difference in the driving performance of the robot.



As explored in the previous design cycle, we could go with a phase-through drivebase. As we are endeavouring for a robot which can clear the supply zone, we believe that this mechanism could be very useful to efficiently help our robot access the supply zone.
To explain again, the phase-through drivebase, we would have a standard tank drivebase. The high supply zone bar is at the perfect height for a wheel to fit underneath. However to get the rest of the robot across the supply zone bar, we need to essentially create a gap in the robot where the bar would pass through. At the back of the robot, we will have a permanent support. Even though this means that the robot cannot fully park with only this mechanism, it means that our robot would not be floating in mid air. However, this would. only provide support to the bar and the front would be stopping forward. Over here, we can use a 3x3 L-shaped piece. This would essentially carry the front of the robot. This piece will be able to support the front of the robot. It would be able to allow the bar to pass through it as it will then turn 90 degrees and rest on the other leg of it.
Even though we originally got this general idea from a YouTube video, as mentioned in our previous design cycle, we have now done our own team enquiry and discussed together how this idea would work in practice.

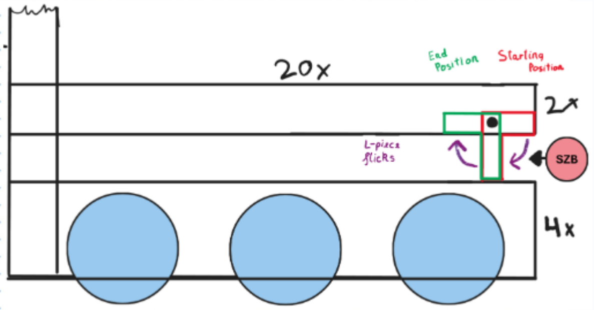
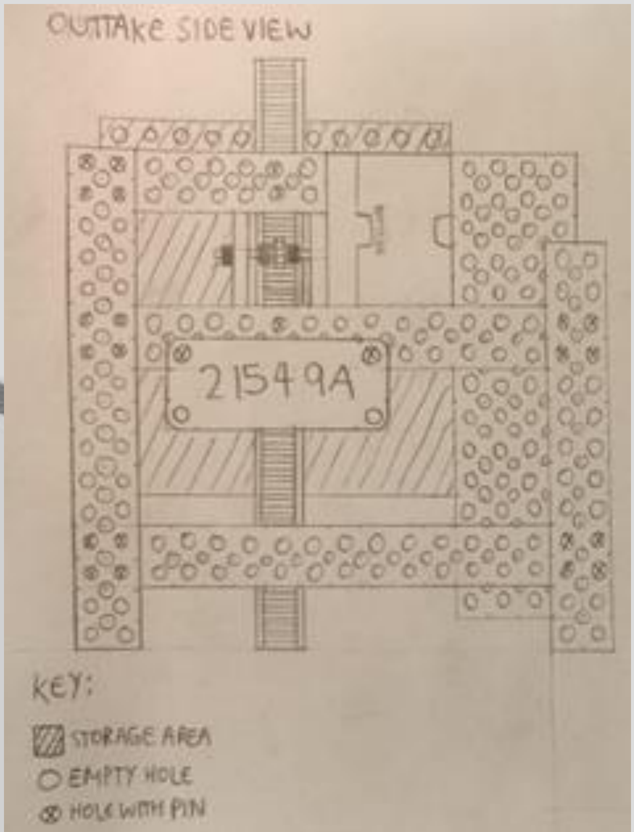
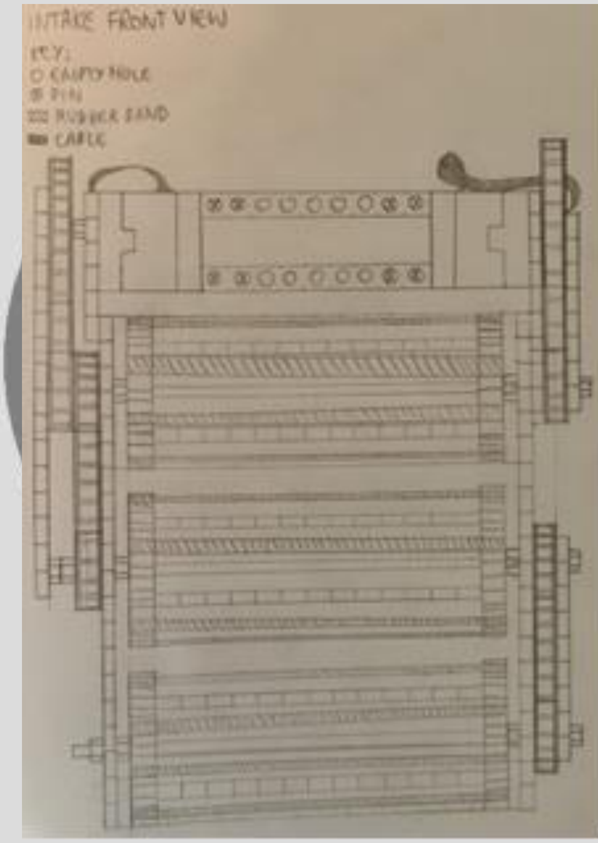
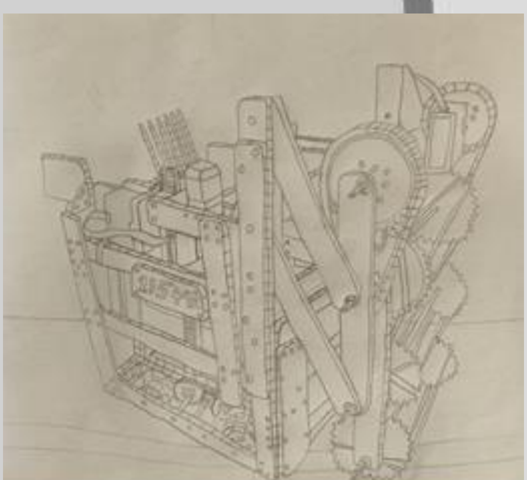


Figure 6: Diagrams created by our team to indicate how different prototype mechanisms would perform

STEP 3: PREPARATION OF THE DATA







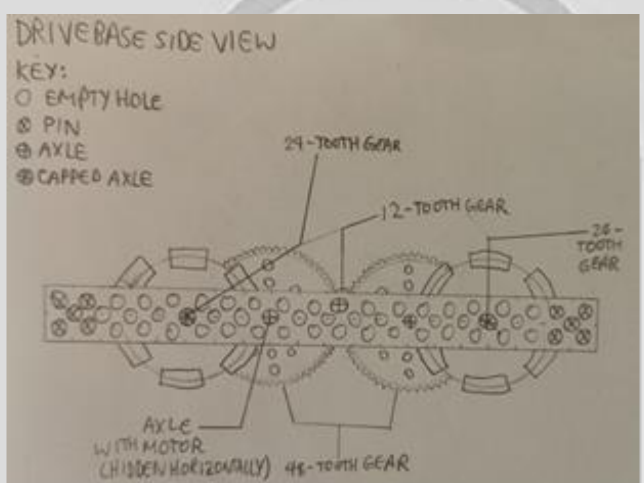


Figure 7: Prototype drawings made to accurately depict mechanisms

STEP 4: MODELLING & EVALUATION

After a model has been completed, data scientists will evaluate the model and its results, and if there are improvements that can be made, the process will be repeated until the model has reached the required standard or been sufficiently improved and can enter the final stage of the design process. This is a process we also follow in our design cycle, where we test an initially completed robot, then go through smaller design cycles to perfect each mechanism of the robot until it is at the level to enter a competition.

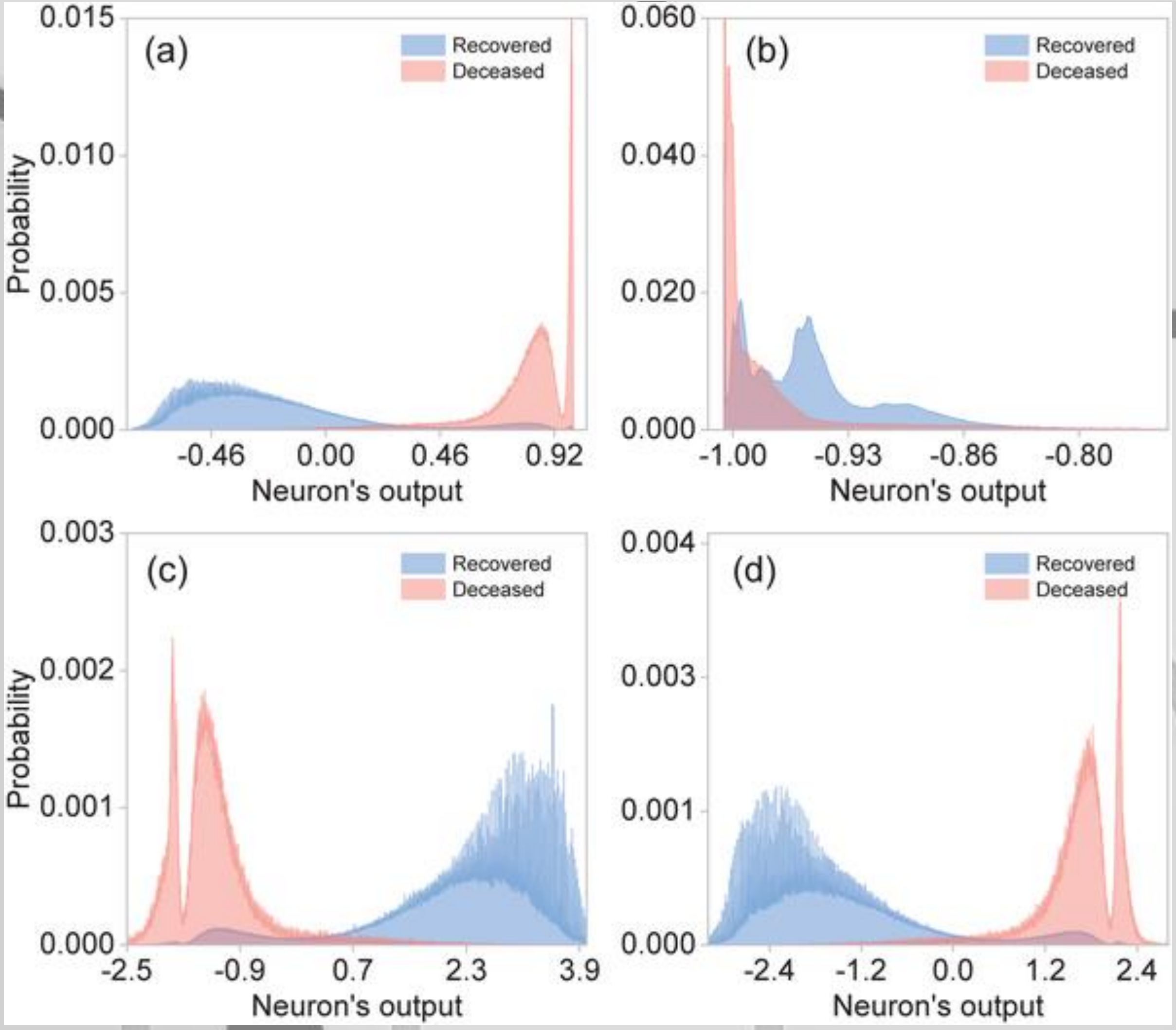
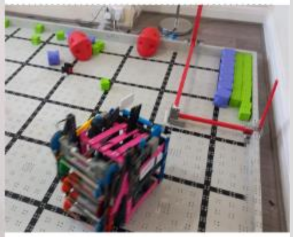


Figure 8: The result of a machine learning model run by Google Cloud as part of the creation of the Covid-19 dataset to recognise the presence of Covid-19 through neural activity

STEP 4: MODELLING & EVALUATION

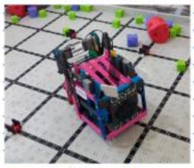
We then decided to test the drivebase speeds and functionality

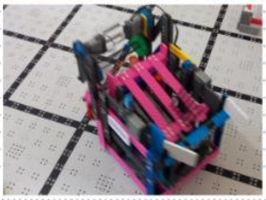


We noticed a problem with the drivebase whilst testing it
This was that sometimes the gears would slip from each other whilst driving fast
This means they wouldn't spin each other and the drivebase would constantly stop working

We fixed this issue by attaching shaft collars to the drive base axles. We had previously been short on shaft collars and been unable to do this. When we did this the gears were kept in place and the issue was resolved

Once this issue was resolved we found out that our drivebase was working really well. We had previously lightened the total weight of the robot and once we fixed the gear issue the drivebase was very fast and could turn easily and quickly. It was also easy to align and very versatile in its uses

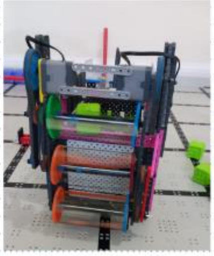




Project Drivebase test

Name Dev Kamesh   Date 13/10/2023   Page 119

After having built our intake and modifying it multiple times, we decided to test it on green cubes to see if any further modifications were needed.



And so, after testing with one green cube at a time, the intake was easily able to intake quickly and rapidly.

We also tested multiple green cubes at once, and they all intaked quite quickly into the storage.

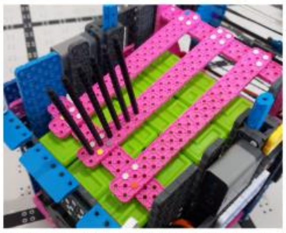
However, in our testing, a problem we did find was the fact that the green cubes would sometime get intake up to the purple section of the storage, losing our 10 points of uniform bonus.

After some iteration, we decided to make the section for greens to enter larger by half a pin, but not too large so that green blocks wouldn't get stuck and go to the top since there was now a larger opening. But, the opening was small enough that the purple cubes still went to the top.
Our testing was useful in that it helped us resolve the issue that green cubes wouldn't go to top where the purples should be.

Project Intake testing (green cubes)

Name Aarav Rathore   Date 12/10/2023   Page 116

We also decided to test how many cubes we could fit in our storage.



We found that our storage had the capacity to fit 24 green cubes in 6 rows of 4 cubes. However, in a real match, the blocks could sometimes jam so achieving this 24 green cubes is very unlikely in a match.

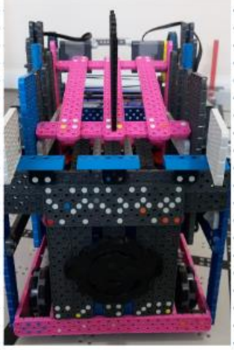
Our top shelf was able to store 6 purple cubes which was also more than enough for a normal game.

And so, after testing, our storage has enough capacity and works well.

Project Storage test

Name Aarav Rathore   Date 12/10/2023   Page 118

We finally tested our rack and pinion outtake mechanism



We wanted to test it to make sure it could flawlessly raise and power itself without any issues in the process. However we noticed an issue
That was that when we raised the outtake it would raise itself too much
That meant the racks would slip out of the brackets
This means we would be unable to lower it back down into a position that we could intake cubes from

We fixed this issue by limiting how high the outtake could go in the code
This meant that the rack could no longer slip out of the brackets, meaning that we could lower the outtake properly

Once we fixed that issue we were satisfied with it fully. Following us having lessened the weight of the outtake it was fast and secure, easily dispensing the cubes without any issues

Project Outtake test

Name Dev Kamesh   Date 13/10/2023   Page 120

STEP 4: MODELLING & EVALUATION

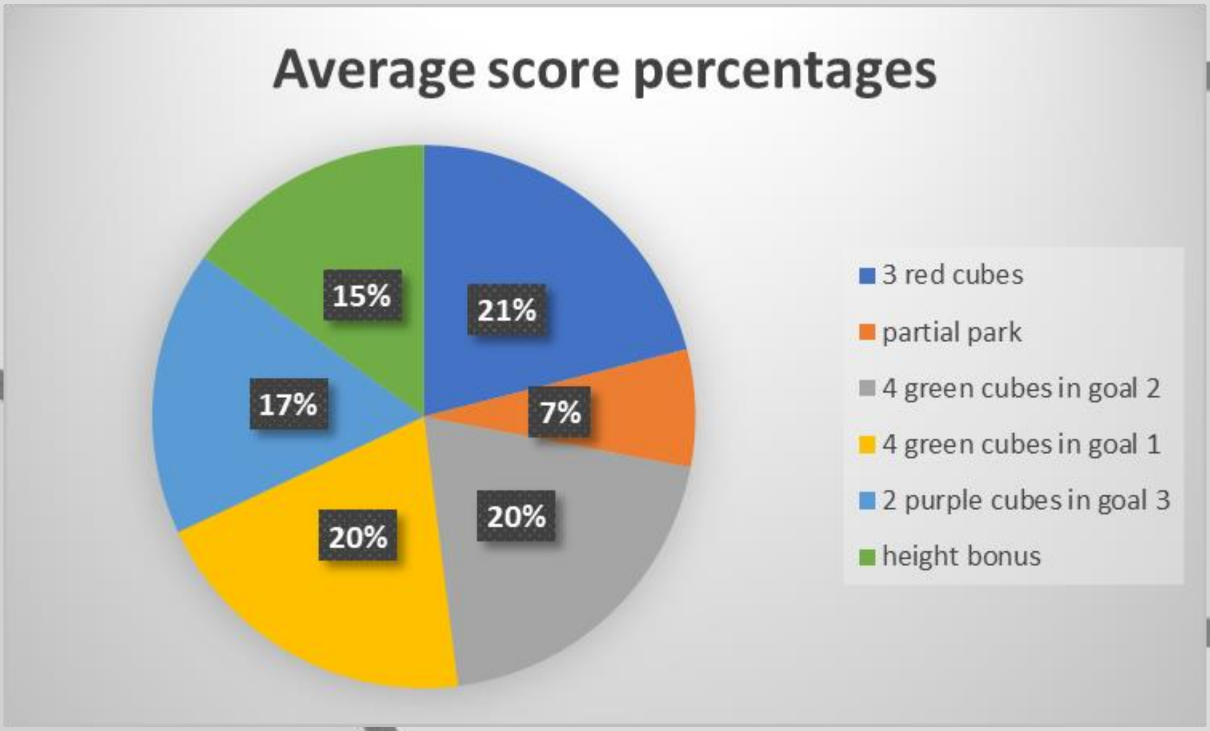


Figure 9: Pages from our Notebook as part of our evaluation stage

STEP 5: DEPLOYMENT

The final stage is deployment, where a finished model and its results are presented to the business where it may be put into use in a real-world application. This presentation usually needs to be simplified or clarified for more effective communication, since data scientists must present to both technical and non-technical audiences. Data scientists are also required to translate their analyses into actionable insights that allow the business to improve its operations and achieve its goals. In this stage, our team begins to do practice runs with the robot and makes sure the drivers are used to the robot in time for a competition. This is the final stage of our design cycle but is still very important as the drivers' skill is a crucial factor to doing well in a competition, however capable the robot itself is.

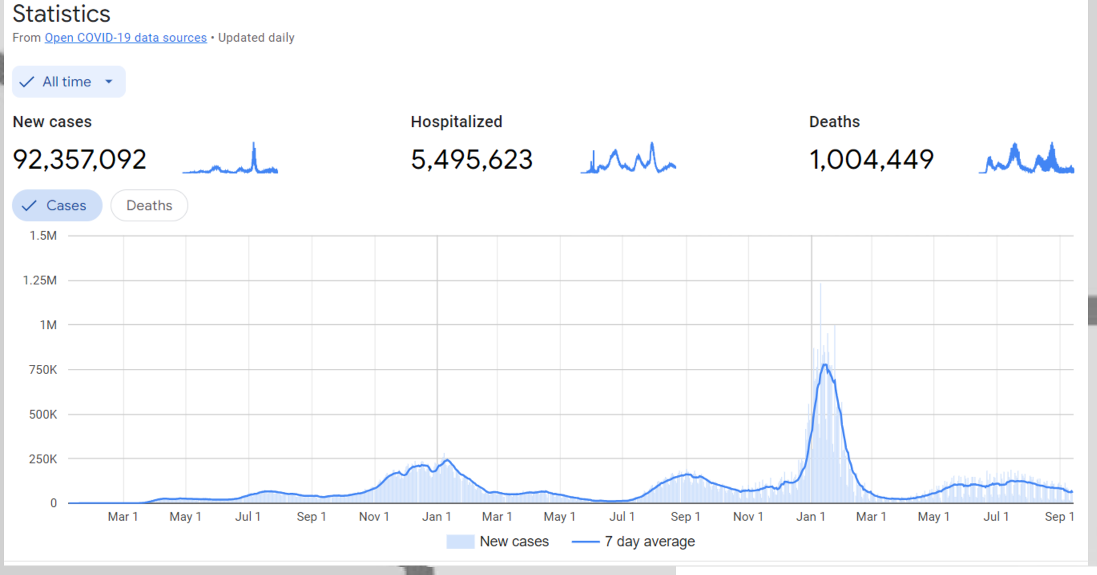
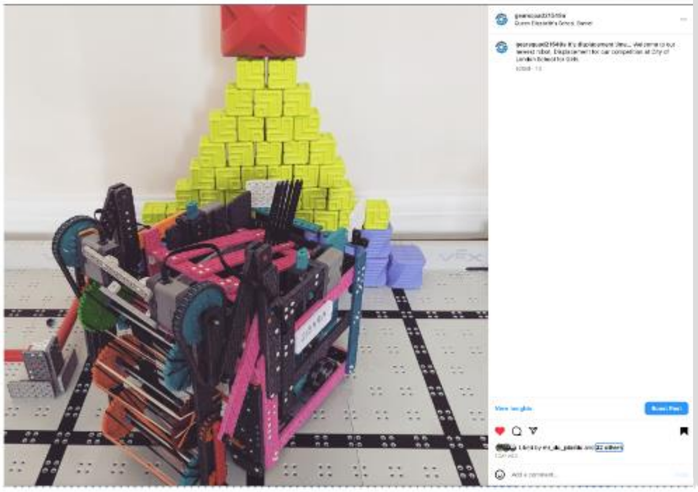


Figure 10: (above) The deployed Covid-19 Dataset after every stage of the design process (below) Our robot in deployment after completing our design cycle

This is the end of our first design cycle of this season before our first competition at City of London School for Girls on Tuesday 31st October 2023.

We have successfully made a robot, which we will be naming "Displacement"



How VEX prepares us for the future.

Participation in VEX Robotics is great preparation for a future career, and not just in STEM. Not only are the skills learnt from VEX useful in careers such as data science and data engineering, but also in other jobs such as technical writing or journalism careers. This is because we do not just practise utilising the engineering design process effectively and working together as a team, but also have good communication skills, which are vital in many other careers outside of STEM that we may choose in the future. In STEM careers, VEX has taught our team the vital procedure of systematically and methodically creating well-researched and effective designs through the design process, which not only ensures that we design to the best of our ability, but also mitigates the effects of time pressure and unexpected problems that often arise from straying from a set design cycle or design sprint. The fundamental principles of design that are learnt through VEX can be applied to just about every engineering situation in the modern world, whether it is application optimisation, building insulation or creating a simulation.

Citations

https://cloud.google.com/
https://cloud.google.com/data-science
https://developers.google.com/learn/topics/datascience
https://www.ibm.com/topics/data-science#:~:text=the%20next%20step-,What%20is%20data%20science%3F,hidden%20in%20an%20organization's%20data.
https://aws.amazon.com/what-is/data-science/
https://towardsdatascience.com/the-data-science-process-a19eb7ebc41b
https://console.cloud.google.com/marketplace/product/bigquery-public-datasets/covid19-public-data-program?hl=en-GB&pli=1
https://www.geeksforgeeks.org/data-science-process/
https://www.applytosupply.digitalmarketplace.service.gov.uk/g-cloud/services/853148900244121
https://www.ibm.com/data-science?utm_content=SRCWW&p1=Search&p4=43700068097012486&p5=e&gclid=Cj0KCQiA2eKtBhDcARIsAEGTG413MYeDADFq2j1vjDoJ7BOTH2KWISIL07_bGX4up2LnfJ8fOGcAF70aAm6FEALw_wcB&gclsrc=aw.ds
https://learn.microsoft.com/en-us/azure/architecture/data-science-process/lifecycle-deployment
https://cloud.google.com/deployment-manager/docs
https://cloud.google.com/blog/topics/developers-practitioners/definitive-guide-databases-google-cloud-part-1-data-modeling-basics

CREDITS

Aarav

I am a designer builder and driver who has participated in VEX IQ for two years. VEX means a lot to me and has helped me learn the skills of collaborating in a team more effectively.

Felix

I am a designer who has participated in VEX IQ for two years. VEX has improved my CAD skills and practical skills.

Avi

I am team captain, designer, builder and driver who has participated in VEX IQ for two years. VEX has taught me some useful life skills like leadership and responsibility.

Dev

I am a designer who has participated in VEX IQ for two years. VEX is important to improve my drawing and technical skills.

Vivaan

I am a programmer and builder who has participated in VEX IQ for two years. VEX has improved my python skills and helped me problem solve.