# Detecting Exoplanets Using AI

By: Gayatri Sakharkar

This slideshow does not cover all information from my logbook here: https://docs.google.com/document/d/1VRiC6cqdC0s0pJS_g7uJ8K475npvgGsHPjTXXnIFjQU/edit

# Question

# Question

How can the application of A.I. (artificial intelligence) improve the accuracy and efficiency of detecting exoplanets through the utilisation of transit photometry data?

# Background Research

# Background Research

What is artificial intelligence?

Artificial intelligence can be described as a branch of computer science that involves simulating the human brain. It is implemented in tasks that require human intelligence. A few of the main functions of artificial intelligence include, problem solving, reasoning, learning, and quick decision making.

At its core, Artificial intelligence is only algorithms with a set of rules. AI systems also have the ability to learn from the repetition of tasks. This how machine learning gets better at doing specific tasks without external interference.

# Background Research

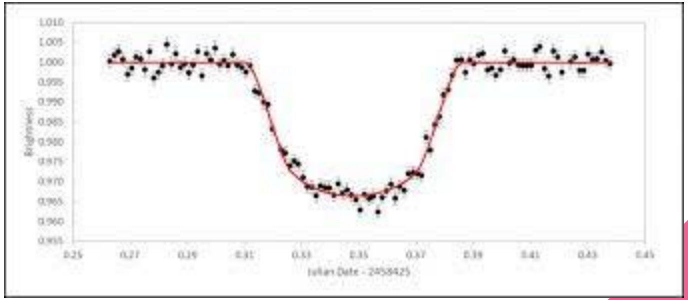
What are exoplanets?

An exoplanet is a planet that orbits a star outside of the solar system.  Some exoplanets, known as rogue planets, do not even orbit stars. Instead, they orbit the galactic center. We know from the Kepler space telescope that there are more exoplanets than stars in the Milky Way Galaxy.

# Background Research

What are scatter plots?

Also known as scatter charts or scatter graphs, scatterplots are graphs that use dots to represent two different quantities. The position of a dot indicates the value of a specific data point. Scatter plots are also used to observe relationships between variables.



# Background Research

For example, in the above scatter plot, the x-axis represents time and y represents the star's brightness. The scatter plot shows the stars' brightness in relation to time.

One common problem when using scatter plots is overplotting. This happens when two or more data points share the same value. A solution to this problem is reducing the number of data points.

# Background Research

What are some exoplanet discovery methods?

1. Radial velocity

Radial velocity is also known as “wobble” this is because sometimes a star will have small movements which make it look like the star is wobbling. This happens because when a planet orbits a star, the gravity of the planet affects the star and makes it move in its own small orbit. This can be observed and help astronomers discover planets beyond the solar system.

# Background research

This method can also be used to identify the mass of the planet or the mass of the planets and how many planets there are because the more there is pulling on the star, the more “wobble” there will be.

The way astronomers are able to observe the wobble of the stars is by the method, Doppler shift. All forms of energy move in waves. The waves change based on how the object producing them moves. For example when you get closer to something that is making sound, it is louder but as you get further away it gets quieter. When sound waves are stretched out it's lower in pitch. Just like that when light waves stretch out it makes the object look more red.

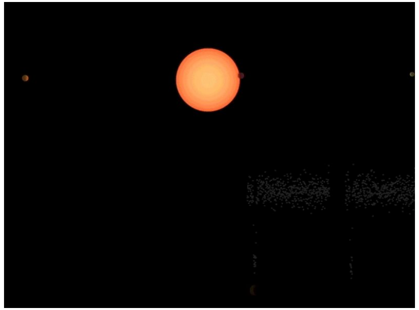
[redacted]

# Background Research

As the star wobbles around, the  light waves change colors and the length of the waves also change which help astronomers find the planet.

# Background Research

2. Transit photometry



What is a Transit?

The majority of the exoplanets discovered were discovered through the transit photometry method. A transit happens when a planet blocks light from the star to the observer, this dims the light very slightly to the observer. When the brightness of the star from the observer's perspective is recorded on a scatter plot, dips can be observed easily.

# Background Research

Transits are not only useful for discovering the exoplanets, but they can tell us a lot about the planet itself. The size of the planet's orbit can be determined by how long it takes for it to orbit its star while the size of the exoplanet can be determined by how much it reduces its star's brightness. It is also possible to learn about the planet's atmosphere by using this method. When the planet is in front of a star, if it has an atmosphere, some light will go through it. This light can later be analyzed to figure out the various elements that make up the atmosphere of the planet. Some space telescopes that have and are collecting data from various stars include Kepler and the Transiting Exoplanet Survey Satellite (also known as TESS).

# Background Research

3. Direct Imaging

Direct imaging as the name suggests, is a method which involves taking a picture of the planet directly. Due to exoplanets being located extremely far away from Earth, it is difficult to discover anything with this method.

One of the main problems with this method is that any light or heat reflected off of the planet is usually drowned out by the planet's host star which makes it difficult to take a picture of it. One solution to this problem are shine blockers. Shine blockers share the same principle as wearing sunglasses, or blocking the sun with your hand to see things which would otherwise not be visible.

# Background Research

Astronomers use two main methods to block light from a star. The first one is called coronagraphy. It uses a device that blocks light from the star before it can reach the telescope's detectors. The second method uses a starshade device. This method blocks light from a star before it even enters the telescope itself.

# Background Research

2. Gravitational Microlensing

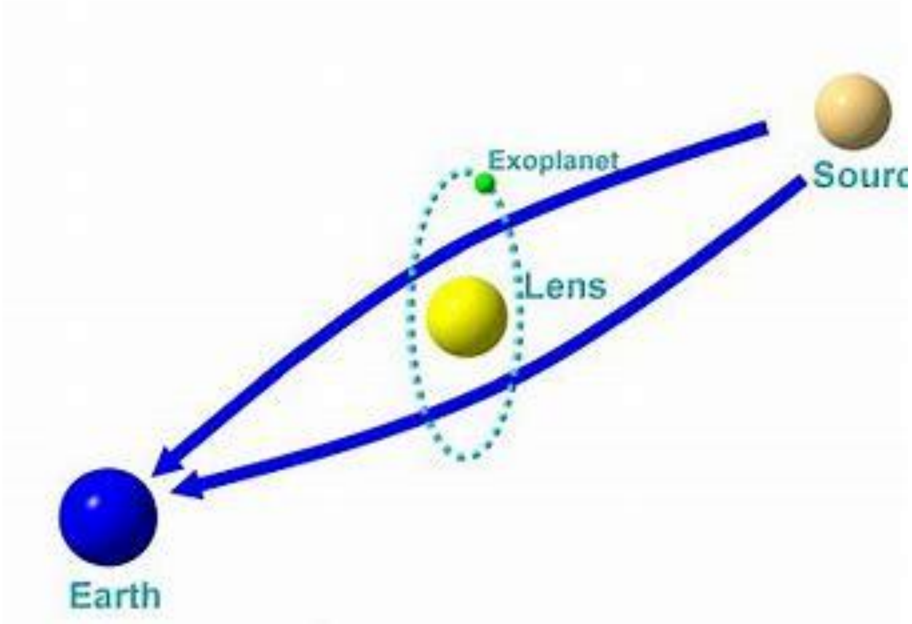
When a huge object, such as a star or planet, affects light, it distorts and changes direction. This is how Albert Einstein redefined gravity as a geometric characteristic of spacetime. This means when a star emits light the effect of gravity changes the direction of light when affected by something by the gravity of a massive object.

# Background Research

For example, when a celestial object passes in front of another object, it appears to look bright. What is happening here is that the object in front bends and magnifies the star further away which makes the further away object look brighter for a short period of time.

If a star passes in front of another star. From our perspective the star further away will look brighter and if the closer star is accompanied by a planet, it will cause the brightness of the star to increase a second time. This helps us figure out if a star has a planet that is traveling with it.

# Background Research



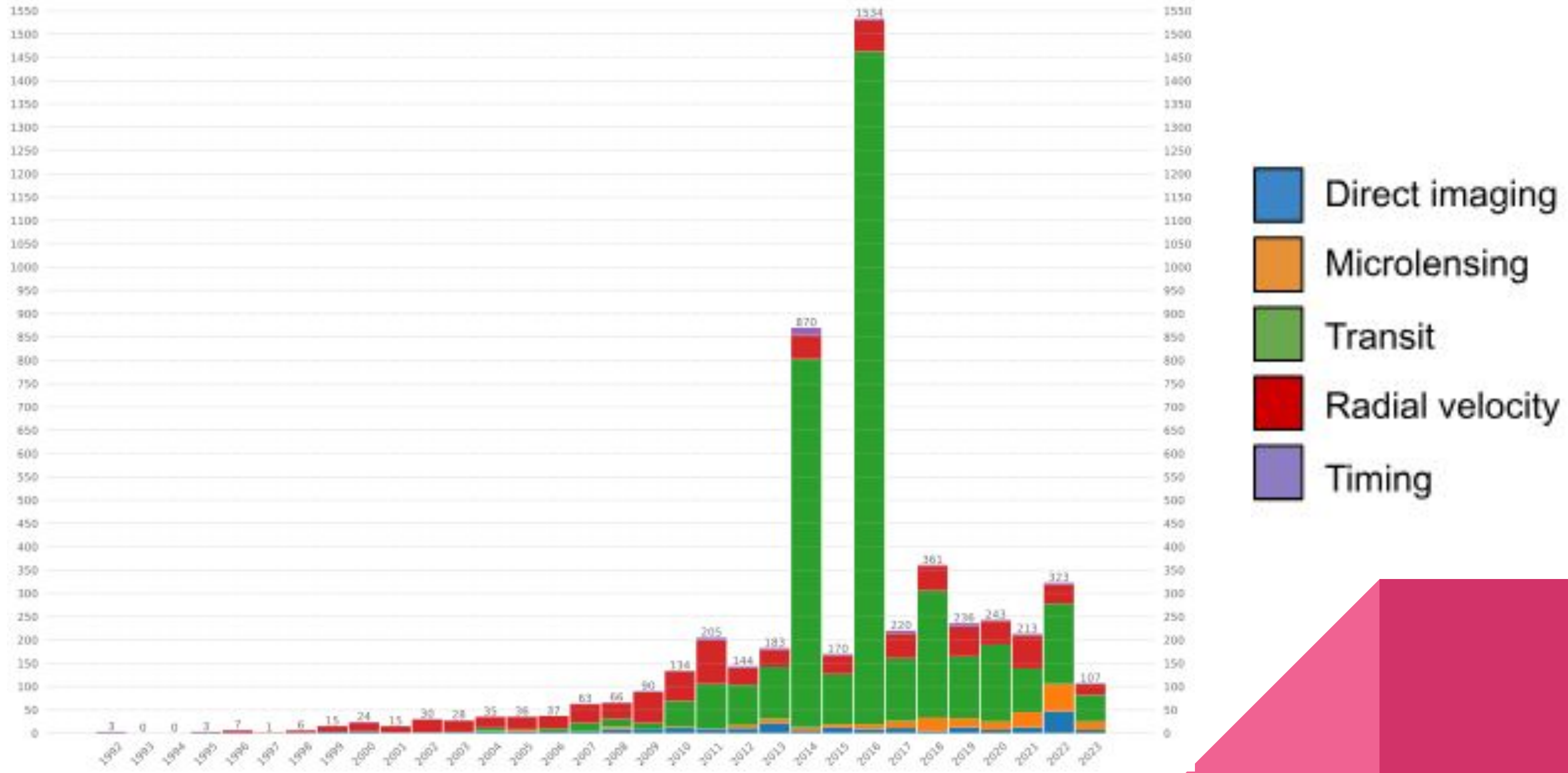
# Background Research

1. Astrometry

Astrometry, like radial velocity, also observes stars’ “wobble”. But here, astronomers do not use the Doppler Shift and instead astronomers are able to see the stars' apparent position in the sky changing.  This method is extremely hard to do and only 3 planets have ever been discovered using it.

Overall, the exoplanet discovering method that has had the most success is the transit photometry method. Here is a graph representing this:

# Background Research



# Background Research

What is the transiting exoplanet survey satellite (TESS)?

The transiting exoplanet survey satellite also known as TESS discovers exoplanets beyond our solar system. It started as a two year mission which was supposed to end on July 4 2020. But it is continuing to discover even more exoplanets as an extended mission. During its original two year mission, it surveyed about 75%. Tess found 66 new exoplanets and 2 100 more candidates that still need to be confirmed.

# Background Research

What is Python and how does it work?

Python is a computer programming language and it is used for a lot of things like building websites, automating tasks, and conducting data analysis. Python is said to be beginner-friendly and easy to learn, making it one of the most used programming languages today. Like most programming languages python uses an input and output. The input in any programming language allows you to interact with the user and the output displays data.

# Background Research

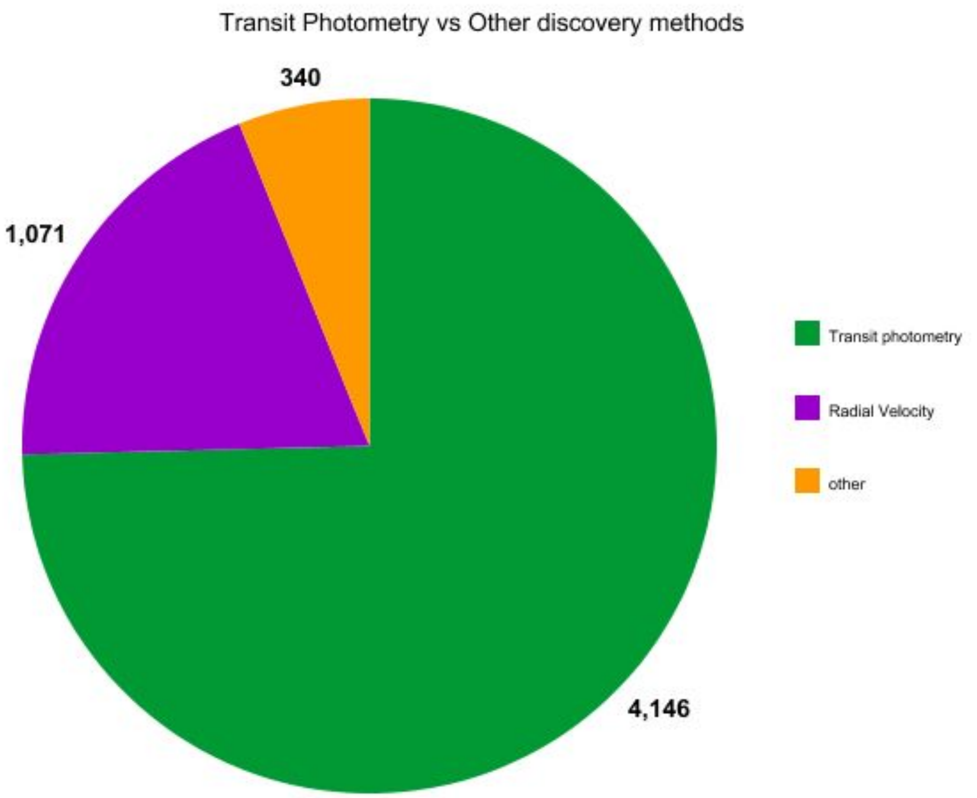
What are traditional methods for discovering exoplanets?

There are many different ways people look for exoplanets. Radial Velocity is one example. This method involves looking for small observable movements in stars caused by gravity of the exoplanet orbiting it. Another method is known as transit photometry. In this method, the brightness of stars is observed and put in scatterplots, scientists then try to observe dips in brightness caused by the planet orbiting it. Direct imaging is another method. Direct imaging involves, as the name suggests, taking pictures of the planet directly.

# Background Research

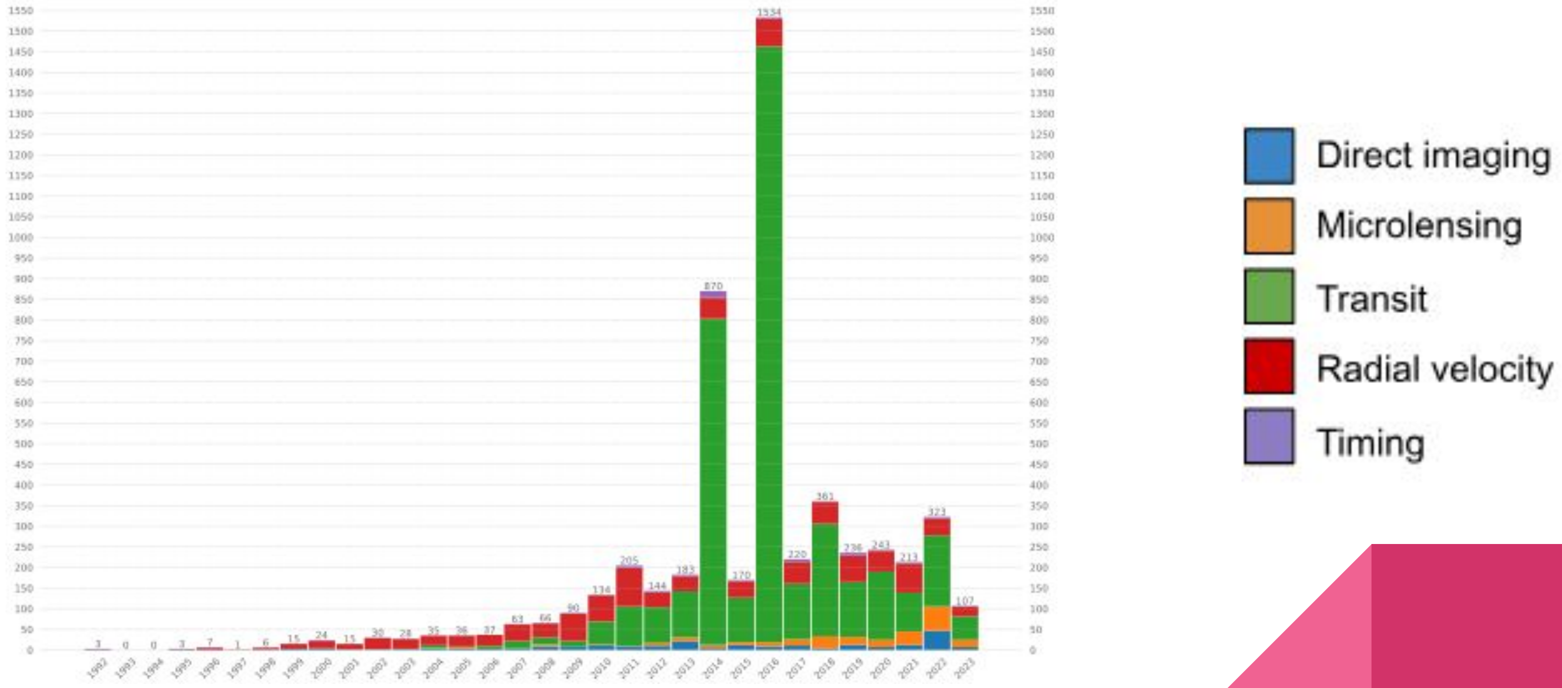
Transit photometry is the best method for finding exoplanets. This is because 4146 have been discovered with the method. This is about 75% of all exoplanets discovered. The transit method can also be compared to other methods to understand how much this method has dominated exoplanet discovery. For example, when compared to radial velocity which has helped discover 1071 exoplanets and is the second most used method for exoplanet discovery only contributes to 19% of all known exoplanets. Here are some graphs showing this information:

# Background Research



12/14/23

# Background Research



# Background Research

For these reasons, transit photometry will be the main focus of this project.

For transit photometry, many traditional methods for exoplanet discovery. Most of these methods involve statistical analysis. For example, Periodicity analysis. Astronomers use this method to detect patterns in data. For example, if a scatterplot shows the brightness of a star, instead of looking solely at how the brightness changes, scientists also look for patterns in the data that repeat.

# Background Research

The brightness of stars does not always stay the same. Instead, it goes through a cycle of getting dimmer and brighter. This cycle happens predictably. Scientists commonly use mathematics to figure out these patterns. For example, Fourier analysis. This is a mathematical technique used in Periodicity analysis. It breaks down complicated signals into simpler parts that repeat which helps scientists see regular patterns clearly.

Another way scientists use traditional methods to discover exoplanets is by using chi-squared tests. This helps scientists determine if observed patterns are real or just caused by random fluctuations in data.

# Background Research

In conclusion, using statistical methods to find exoplanets can be a great way to discover new planets. Though it is crucial to note that using these methods may be complicated and time consuming. This is why AI may be useful in making exoplanet discoveries more simple and efficient.

# Background Research

How is AI being used to discover exoplanets?

Traditional methods that involve statistical analysis can be amazing methods to discover exoplanets. One big problem with this however, is the complexity of data collected from light curves. takes a long time. Even up to years to discover a single planet.

AI is excellent at detecting patterns, therefore AI can be used to discover dips and predict future patterns in light curves as well. This can possibly be done much more efficiently and accurately than traditional methods.

# Background Research

Machine learning for example, has countless applications in exoplanet detection in transit photometry data. Firstly, it can recognise patterns in transit photometry data. These algorithms are able to learn patterns and identify dips in brightness caused by planets orbiting their host stars. This can greatly reduce the time and effort required to manually analyse data.

Machine learning can also be used for signal processing. Transit photometry data can be complicated. Machine learning models may be trained to differentiate between transit signals and other variability in the light curve. This can help scientists because the process of filtering random fluctuations greatly improves the accuracy of the detection of exoplanets.

# Background Research

Deep learning can also be used. Deep learning is a subunit of machine learning which is inspired by the human brain. It can be used to discover exoplanets because of its ability to automatically find exoplanets by looking at light curves. This can especially be useful for large-scale surveys like TESS (the transiting exoplanet survey satellite).

In conclusion, AI can be excellent at detecting exoplanets because of its ability to automatically identify dips in scatterplots. AI has the potential not only to reduce the time it takes to detect exoplanets but also to enhance the accuracy of the process.

# Hypothesis

# Hypothesis

If AI is used to aid in exoplanet discovery, then the efficiency and accuracy of detecting exoplanets will be improved because AI is extremely effective at identifying patterns. This is significant because transit photometry, which is represented on scatter plots, has been proven to be the most successful method to discover exoplanets. Instead of manually detecting dips in these scatter plots, AI can automate the task and make it much simpler.

# Procedure

# Materials

- Printer
- Computer with internet access
- Printed “Star A” scatter plot (x10)
- Printed “Star B” scatterplot (x10)
- Printed “Star C” scatterplot (x10)
- Printed “Star D” scatterplot (x10)
- Printed “Star E” scatterplot (x10)
- Printed “Star F” scatterplot (x10)
- Pen
- Stopwatch
- Calculator

# Procedure

1. Review literature
    a. Gather information about how traditional methods are being used to discover exoplanets
    b. Gather information about how AI methods are being used to discover exoplanets
    c. Gather information that compares AI and traditional methods
    d. Understand various concepts related to the project including:
        - Scatterplots
        - Transit photometry
        - Deep learning
        - Supervised ML algorithms
        - Data regression

# Procedure

2. Finding a program
    a. Choose an AI algorithm that can detect dips in scatterplots
    b. This program must Identify dips in scatterplots (which can possibly be an exoplanet) by using a line of perfect fit to find underlying amounts (the lowest part of the line of perfect fit will determine the maximum value to be considered a dip.
    c. Use program found in observations to replicate this experiment
3. Create 6 different scatter plots
    a. Generate random data points that represent star brightness data from ChatGPT (this will be the simulated star brightness data used in the experiments)
    b. Apply the algorithms on the scatterplots to identify the dips
    c. Record all observations including difficulties, and how many dips were identified correctly.

# Procedure

4. Run the experiment on human test subjects
    a. Explain to 10 different human test subjects' what transit photometry is and how it works
    b. Explain to 10 different human test subjects' what are scatter plots
    c. Test subjects manually detect dips in scatterplots (the same scatter plots will be used from step 2)
    d. Write down what test subject is being tested on on each graph
    e. Have each test subject analyse each of the six scatterplots
    f. Record all observations including difficulties in identifying dips, how many dips were identified correctly and how much time it had taken to analyse the scatter plot
5. Run the experiment on the program
    a. Copy and paste star brightness data in the respective area
    b. Run program
    c. Record how many dips were identified correctly and incorrectly
    d. Record how many dips were not identified

# Procedure

6. Analyse quantitative data
    a. Using the following formula determine overall accuracy of the test subject or algorithm: point that are actually dips divided by total points identified as dips by test subject or algorithm
    b. Repeat step 7.a for every graph analysed by human test subject as well as the test subjects themselves
    c. Repeat step 7.a for every graph analysed by the program
    d. Record all data
    e. Use the following data to complete analysis
        - Overall accuracy of program and humans
        - Incorrectly identified dips
        - Correctly identified dips
        - Unidentified dips
        - Time taken to analyse graph
    f. Make bar graphs with the data to show information

# Procedure

7. Conclude how and if AI is effective in detecting exoplanets
    a. Use qualitative and quantitative data
    b. Explain how AI is beneficial
    c. Discuss potential applications of AI

# Observations

# Research

## Part 2 - Quantitative data

## Simulated star data:

Star A -

- Hours Observed: [10, 15, 8, 20, 25, 12, 18, 30, 35, 28, 22, 15, 20, 25, 18, 22, 12, 28, 15, 32]
- Brightness: [5, 7, 4, 8, 2, 6, 3, 9, 12, 8, 5, 3, 7, 10, 4, 6, 2, 8, 5, 11]

Star B -

- Brightness: [12, 10, 8, 15, 5, 18, 6, 13, 9, 14, 8, 5, 11, 16, 7, 10, 4, 12, 15, 6]
- Hours Observed: [8, 12, 10, 18, 22, 15, 20, 32, 28, 35, 25, 18, 30, 22, 28, 15, 10, 32, 18, 25]

# Research

Star C -

- Brightness [3, 5, 2, 6, 10, 4, 8, 3, 7, 5, 9, 2, 6, 11, 4, 7, 3, 8, 5, 10]
- Hours Observed: [25, 20, 30, 15, 10, 22, 18, 35, 28, 32, 22, 15, 25, 30, 18, 20, 28, 12, 22, 35]

Star D -

- Brightness: [8, 10, 6, 12, 4, 9, 5, 11, 7, 14, 9, 6, 10, 15, 8, 12, 4, 13, 6, 11]
- Hours Observed: [18, 15, 20, 25, 30, 22, 10, 28, 32, 35, 22, 15, 20, 25, 18, 30, 12, 28, 15, 25]

Star E -

- Brightness: [6, 4, 7, 5, 8, 3, 9, 5, 10, 4, 6, 2, 8, 11, 7, 3, 6, 9, 5, 12]
- Hours Observed: [12, 18, 15, 10, 20, 25, 30, 22, 35, 28, 18, 15, 25, 30, 12, 20, 28, 22, 15, 32]

# Research

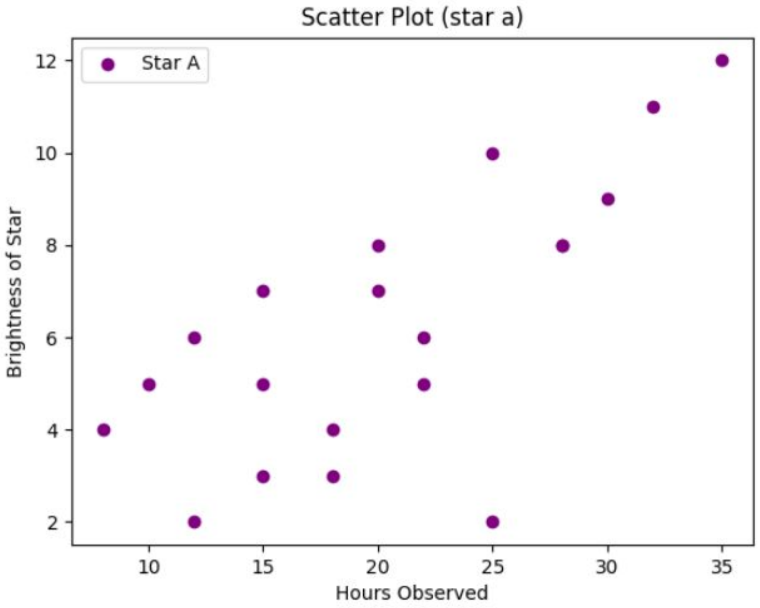
Star F -

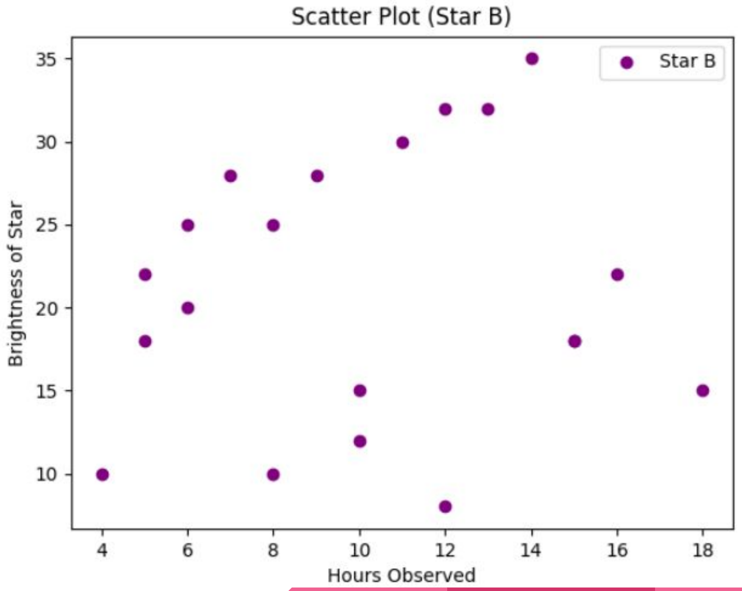
- Brightness: [8, 10, 6, 12, 4, 9, 5, 11, 7, 14, 9, 6, 10, 15, 8, 12, 4, 13, 6, 11]
- Hours Observed: [18, 15, 20, 25, 30, 22, 10, 28, 32, 35, 22, 15, 20, 25, 18, 30, 12, 28, 15, 25]

12/14/23

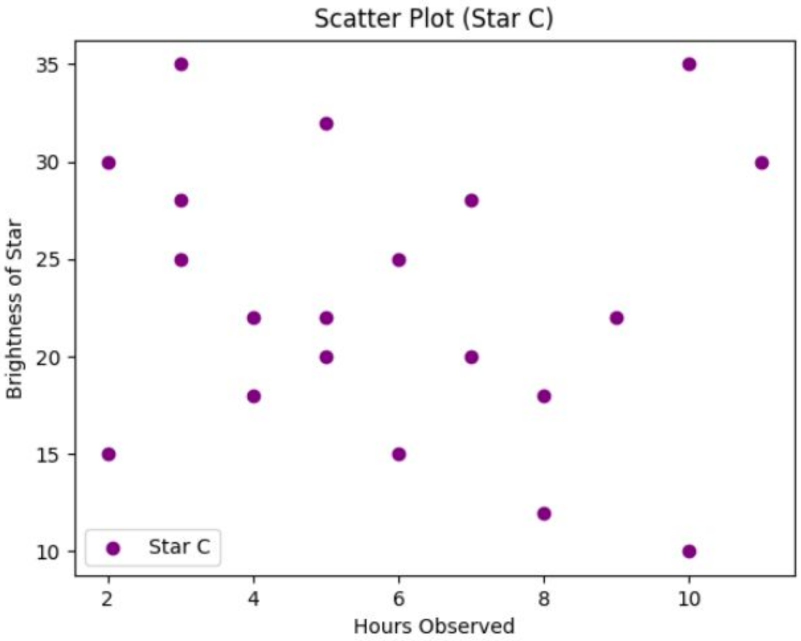
# Research

Star data put into scatterplots

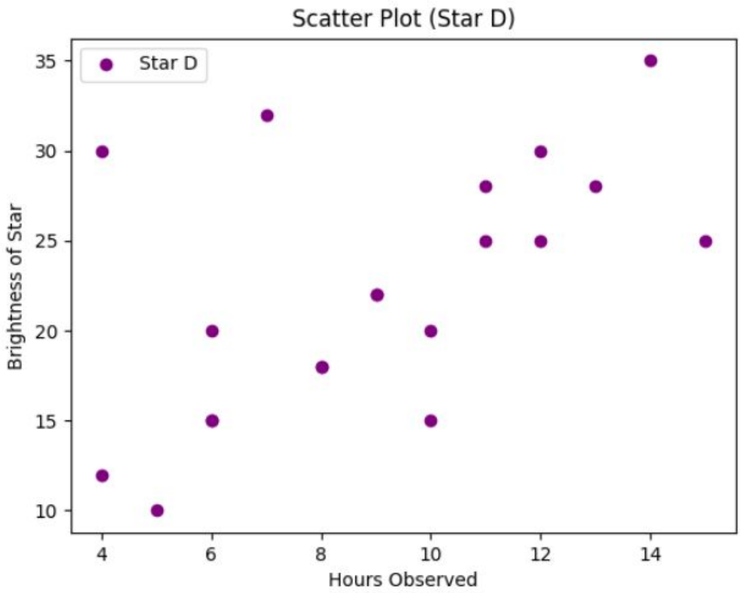


# Research

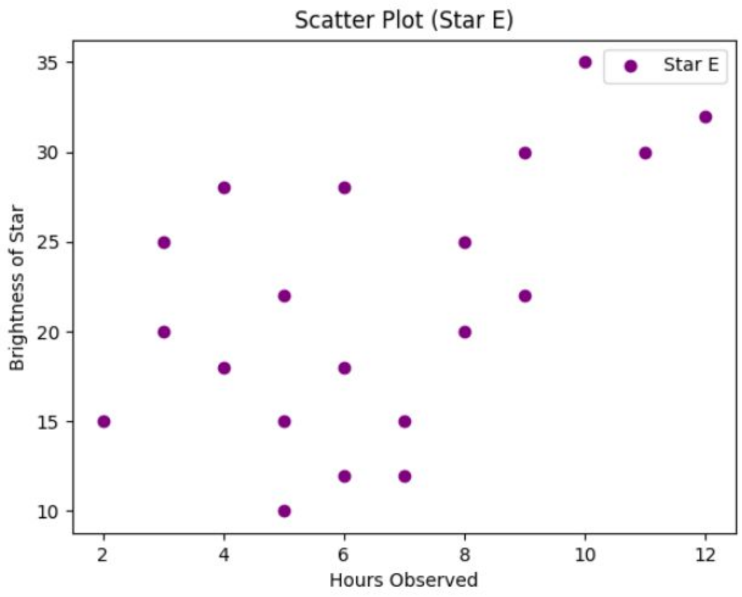


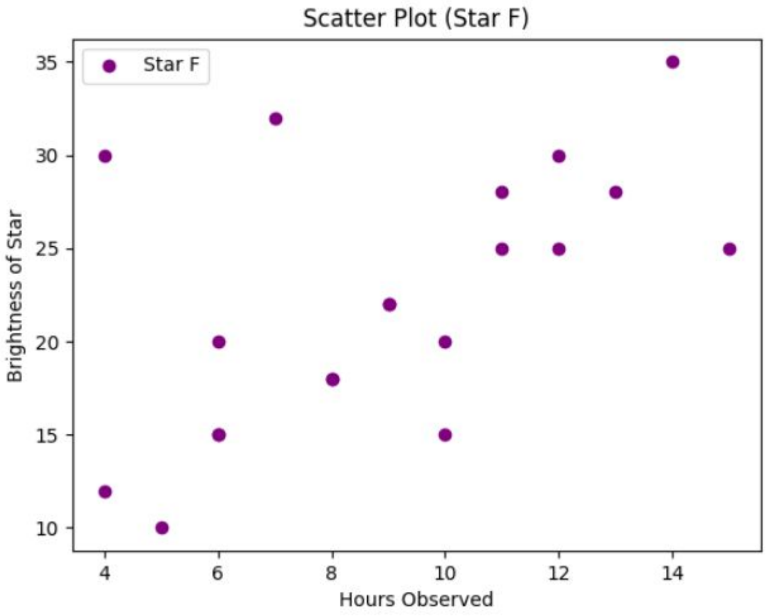
Scatter Plot (Star C)
Star C
Brightness of Star
Hours Observed
35
30
25
20
15
10
2
4
6
8
10

Scatter Plot (Star D)
Star D
Brightness of Star
Hours Observed
35
30
25
20
15
10
4
6
8
10
12
14

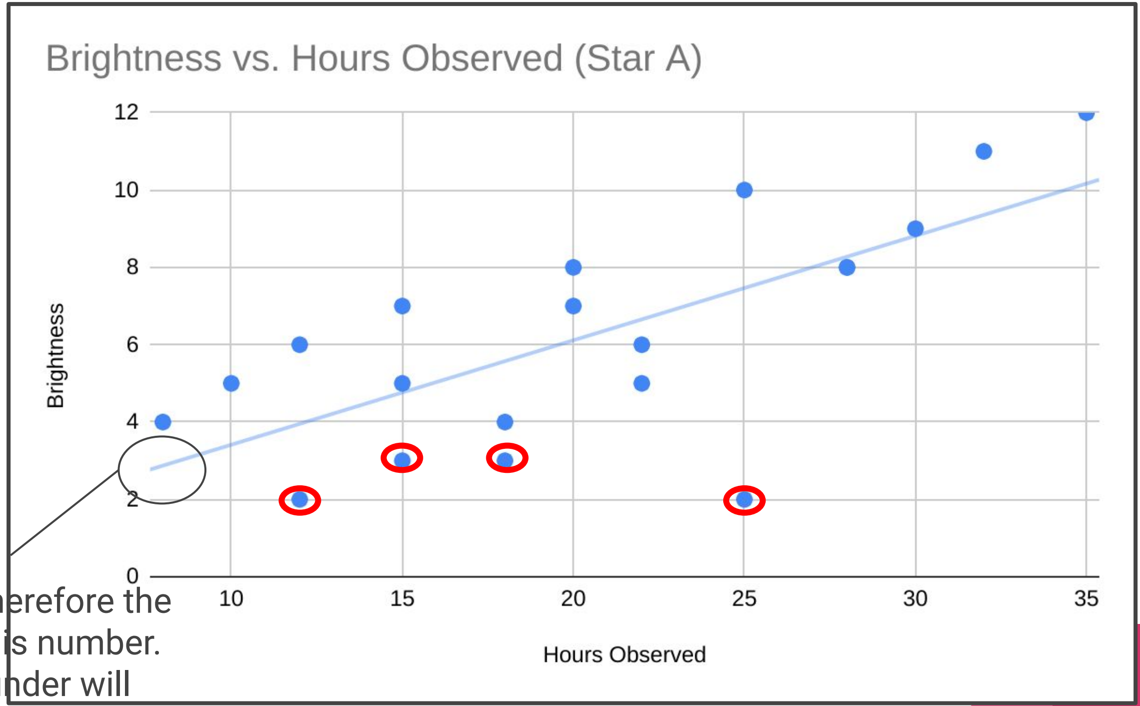
# Research


Scatter Plot (Star E)
Star E
Brightness of Star
Hours Observed
35
30
25
20
15
10
2
4
6
8
10
12


Scatter Plot (Star F)
Star F
Brightness of Star
Hours Observed
35
30
25
20
15
10
4
6
8
10
12
14

# Research

This scatter plot includes the line of best fit




The lowest point is 3, therefore the underlying amount is this number. Anything that is 3 and under will be considered a dip.

# Research

<u>Testing Observations</u>

12/15/23 - started testing. Only did one graph on test subjects 1 and 2

- It takes a long time to explain the concept of exoplanets and scatter plots to test subjects. This can be skipped in AI
- Both test subject one and two circled the points (15, 3) and (18,3)
- Both took thirty seconds to complete star a

# Research

12/16/23 - continued to test. Finished all six graphs on test subjects one and two

- For test subject one, scatter plots for Star A and E were the most difficult to analyse as the points were too close to each other.
- For test subject one, It was hard to determine if the dips were due to an exoplanet or other factors.
- Finished testing for test subjects 1 and 2
- As test subject 1 and 2 continue to analyse graphs the time it took reduced to about 17 seconds each
- For test subject two, the first graph was also difficult because they did not understand the concept.
- For test subject two, they understood the concept by visualising the star and how

# Research

- the exoplanet reduces the brightness of the star. This means that it should be explained how exoplanets reduce star brightness from the perspective of Earth. This will ensure clarity about the topic.

12/17/23 - completed more testing on human participants. Had three participants tested at the end of the day

- For test subject three, they mentioned that they liked how I explained the tests
- For test subject three, they mentioned that that they found the test interesting
- For test subject three and four, they mentioned how they found it easier as they analysed more scatterplots
- For test subject three and four, both found star A the hardest because it was the first one and it was hard to understand
- For test subject five, they used a method which involved looking for an average layer of dots. Any dot that was under this average layer, was considered a dip.
- For test subject five, they found star c the most difficult because it was not very easy to identify the average layer.
- For test subject five, they mentioned how I should have explained the tests better for example, explain transit photometry in detail, how data will be used and explain exoplanets better

# Research

12/19/23 - completed all of human testing

- For test subject six, they found all scatter plots were the same difficulty
- For test subject six, mentioned how it was not easy to identify recurring patterns in the graphs
- Planning to test more human subjects
- Test subject six did not have a pattern in the time as in their time was relatively similar the entire time and did not reduce
- The sixth test subject's time and accuracy appear to be significantly greater than the other participants.

# Research

12/22/23 - decided to continue human testing

- For test subject seven, they thought all were the same difficulty
- For test subject seven, they thought I explained it nicely but they also had done research prior to the experiment
- For test subject the accuracy looked like it was excellent
- For test subject eight, they mentioned how it is pointless for humans to do something like this and it is better to use a program to identify dips in scatterplots
- For test subject eight, the experiment was simply confusing but not hard because it is difficult to get an exact line of perfect fit
- For test subject eight, they mentioned how it was just based on opinion (was is a dip and what is not)

# Research

12/23/23 - completed testing on ten participants

- For test subject nine, there was no graph that was the easiest or hardest they were all the same difficulty and they were all hard
- For test subject nine, they mentioned how all the scatterplots looked the same
- For test subject ten, star F was easiest because all dips were mostly together
- For test subject ten, star c was the hardest because the points were more scattered
- For test subject ten, they thought my explanation was good
- It is interesting to see how the data varies so much
- A question that may be asked is. Why are some people better than others at identifying dips in scatterplots? Is it because of lack of understanding or skill?
- The fact that it is possible that my explanation is not good enough can either be a source of error or another way AI is better than humans. This is because AI just

# Research

- needs programming while it can take years to educate just one person about statistical analysis until they are ready to properly identify the dips.

12/24/23 - Analysed scatterplots analysed by hy humans
12/25/23 - Analysed scatterplots analysed by humans. Started testing on algorithms. Encountered a problem.

- The AI algorithm was tested on star A but it was not accurate and all the points were not detected (the accuracy was 50%)
- After the realisation that it was not working, continued testing the algorithm on other scatter plots to see if the same problem would be encountered
- The other scatter plots were 100% accurate
- Decided to check if the points that were thought to be dips were actually dips but they were not dips
- It is now necessary to change all data about star a

# Research

Human participants: Time to detect all Dips (minutes)

| | Star A | Star B | Star C | Star D | Star E | Star F |
|---|---|---|---|---|---|---|
| Test subject 1 | 0.29.31 | 0.26.60 | 0.25.68 | 0.18.10 | 0.18 | 0.18.61 |
| Test subject 2 | 0.29.06 | 0.23.65 | 0.17.49 | 0.16.49 | 0.16.32 | 0.20.83 |
| Test subject 3 | 2.27.0 | 0.29.68 | 0.21.34 | 0.23.63 | 0.29.84 | 0.46.0 |
| Test subject 4 | 3.48.0 | 0.57.64 | 0.44.94 | 0.38.52 | 0.36.54 | 1.25.0 |
| Test subject 5 | 0.57.66 | 0.9.8 | 0.29.5 | 0.26.9 | 0.57.4 | 0.23.2 |
| Test subject 6 | 0.13.82 | 0.9.42 | 0.15.39 | 0.9.36 | 0.8.41 | 0.6.86 |

# Research

| | | | | | | |
|---|---|---|---|---|---|---|
| Test subject 7 | 23.87 | 13.08 | 17.10 | 18.17 | 9.02 | 10.92 |
| Test subject 8 | 20.01 | 22.82 | 25.10 | 5.59 | 17.28 | 19.55 |
| Test subject 9 | 1.03.48 | 0.37.50 | 0.40.30 | 0.29.52 | 0.21.05 | 0.16.48 |
| Test subject 10 | 1.08.20 | 0.16.04 | 0.14.08 | 0.19.18 | 13 | 10.28 |

# Research

| | A | B | C | D | E | F |
|---|---|---|---|---|---|---|
| Average amount of time taken to detect all dips (minutes) | 1.002 | 0.354 | 0.251 | 0.342 | 0.379 | 0.361 |

# Research

*Points identified correctly*

| | Star A | Star B | Star C | Star D | Star E | Star F |
|---|---|---|---|---|---|---|
| Test subject 1 | 2 | 6 | 6 | 4 | 6 | 4 |
| Test subject 2 | 0 | 4 | 3 | 3 | 1 | 3 |
| Test subject 3 | 2 | 6 | 4 | 2 | 5 | 2 |
| Test subject 4 | 1 | 0 | 1 | 0 | 1 | 1 |
| Test subject 5 | 1 | 7 | 7 | 2 | 5 | 2 |

# Research

| | | | | | | |
|---|---|---|---|---|---|---|
| Test subject 6 | 2 | 4 | 2 | 2 | 3 | 2 |
| Test subject 7 | 2 | 4 | 3 | 2 | 2 | 2 |
| Test subject 8 | 1 | 6 | 1 | 1 | 4 | 1 |
| Test subject 9 | 0 | 3 | 3 | 2 | 2 | 3 |
| Test subject 10 | 1 | 8 | 8 | 4 | 6 | 4 |

# Research

| | Star A | Star B | Star C | Star D | Star E | Star F |
|---|---|---|---|---|---|---|
| Average amount of dips unidentified by human participants | 1.2 | 4.8 | 2.2 | 3.8 | 3.5 | 2.4 |

# Research

Points identified incorrectly

| | Star A | Star B | Star C | Star D | Star E | Star F |
|---|---|---|---|---|---|---|
| Test subject 1 | 3 | 0 | 0 | 1 | 2 | 1 |
| Test subject 2 | 7 | 0 | 1 | 2 | 2 | 1 |
| Test subject 3 | 9 | 3 | 2 | 6 | 6 | 6 |
| Test subject 4 | 1 | 2 | 0 | 2 | 1 | 1 |
| Test subject 5 | 4 | 0 | 0 | 3 | 2 | 4 |

# Research

| | | | | | | |
|---|---|---|---|---|---|---|
| Test subject 6 | 2 | 0 | 0 | 0 | 0 | 0 |
| Test subject 7 | 2 | 0 | 0 | 0 | 0 | 1 |
| Test subject 8 | 2 | 1 | 4 | 1 | 0 | 3 |
| Test subject 9 | 7 | 3 | 3 | 4 | 2 | 3 |
| Test subject 10 | 9 | 2 | 4 | 4 | 6 | 4 |

# Research

| | Star A | Star B | Star C | Star D | Star E | Star F |
|---|---|---|---|---|---|---|
| Average Amount of Dips Identified correctly by Human Participants | 4.6 | 1.1 | 1.4 | 2.3 | 2.1 | 2.4 |

# Research

Points that were unidentified

| | Star A | Star B | Star C | Star D | Star E | Star F |
|---|---|---|---|---|---|---|
| Test subject 1 | 0 | 2 | 2 | 0 | 0 | 0 |
| Test subject 2 | 2 | 4 | 5 | 1 | 5 | 1 |
| Test subject 3 | 0 | 2 | 3 | 2 | 1 | 2 |
| Test subject 4 | 1 | 8 | 7 | 4 | 4 | 3 |
| Test subject 5 | 1 | 1 | 1 | 2 | 1 | 2 |

# Research

| | | | | | | |
|---|---|---|---|---|---|---|
| Test subject 6 | 0 | 4 | 6 | 2 | 3 | 2 |
| Test subject 7 | 0 | 4 | 5 | 2 | 4 | 2 |
| Test subject 8 | 1 | 2 | 7 | 3 | 2 | 3 |
| Test subject 9 | 2 | 6 | 5 | 2 | 4 | 1 |
| Test subject 10 | 1 | 0 | 0 | 0 | 0 | 0 |

# Research

| | Star A | Star B | Star C | Star D | Star E | Star F |
|---|---|---|---|---|---|---|
| Average amount of dips unidentified by human participants | 0.8 | 3.3 | 4.1 | 1.8 | 2.4 | 1.6 |

# Research

Overall accuracy (percents)

| | Star A | Star B | Star C | Star D | Star E | Star F |
|---|---|---|---|---|---|---|
| Test subject 1 | 40 | 75 | 75 | 80 | 75 | 80 |
| Test subject 2 | 0 | 50 | 33 | 50 | 13 | 60 |
| Test subject 3 | 18 | 55 | 44 | 20 | 42 | 20 |
| Test subject 4 | 33 | 0 | 13 | 0 | 17 | 20 |
| Test subject 5 | 17 | 88 | 88 | 29 | 63 | 25 |

# Research

| | | | | | | |
|---|---|---|---|---|---|---|
| Test subject 6 | 50 | 50 | 25 | 50 | 50 | 50 |
| Test subject 7 | 50 | 50 | 38 | 50 | 33 | 40 |
| Test subject 8 | 25 | 67 | 8 | 20 | 67 | 15 |
| Test subject 9 | 0 | 25 | 27 | 25 | 25 | 43 |
| Test subject 10 | 9 | 80 | 75 | 50 | 50 | 50 |

# Research

| | Star A | Star B | Star C | Star D | Star E | Star F |
|---|---|---|---|---|---|---|
| Average accuracy of human participants (percentage) | 24.2 | 54 | 42.6 | 37.4 | 43.5 | 40.3 |

# Research

AI: Accuracy

The AI algorithm completes analysis within milliseconds, therefore, there is no data for how long the algorithm took to identify the dips in the scatter plots.

Points identified correctly

| | Star A | Star B | Star C | Star D | Star E | Star F |
|---|---|---|---|---|---|---|
| AI algorithm | 2 | 8 | 8 | 4 | 6 | 4 |

# Research

Points identified incorrectly

| | Star A | Star B | Star C | Star D | Star E | Star F |
|---|---|---|---|---|---|---|
| AI algorithm | 0 | 0 | 0 | 0 | 0 | 0 |

# Research

Points unidentified

| | Star A | Star B | Star C | Star D | Star E | Star F |
|---|---|---|---|---|---|---|
| AI algorithm | 0 | 0 | 0 | 0 | 0 | 0 |

# Research

Overall Accuracy (percents)

| | Star A | Star B | Star C | Star D | Star E | Star F |
|---|---|---|---|---|---|---|
| AI algorithm | 100 | 100 | 100 | 100 | 100 | 100 |

# Analysis

# Analysis

The results of the experimental part of this project prove that AI is much better at identifying dips in scatterplots. This is because it took an average of 44 seconds for the human participants to identify the dips in each scatterplots. This is 44% more than AI.

The results also show that human participants could not correctly identify as many dips as AI. This is because on star A, an average of 1.2 dips were identified correctly by human participants while the AI algorithm was able to detect all 2 dips correctly. This is about 67% more.

For star B, human participants were able to identify an average of 4.8 dips correctly. The AI algorithm was able to detect 8 correctly. This is about 67% more.

# Analysis

For star C, human participants were able to identify an average of 2.2 dips correctly. The AI algorithm was able to detect 8 correctly. This is about 264% more.

For star D, human participants were able to identify an average of 3.8 dips correctly. The AI algorithm was able to detect 4 correctly. This is about 5% more.

For star E, human participants were able to identify an average of 3.5 dips correctly. The AI algorithm was able to detect 4 correctly. This is about 14% more.

For star F, human participants were able to identify an average of 2.4 dips correctly. The AI algorithm was able to detect 4 correctly. This is about 67% more.

# Analysis

The findings demonstrated that AI is far more accurate than humans in many aspects. This is because AI is not only capable of accurately detecting dips in scatter plots, but it is also capable of avoiding wrongly recognizing dips and missing any points that represent dips. This is demonstrated by the fact that in all scatter plots, it is zero in the areas that indicate how many dips were incorrectly identified and missed.

Overall accuracy was also measured. This was done in percentages by dividing the amount of correctly identified points to the sum of correctly identified points, incorrectly identified points, and missed points.

For star A, the average accuracy was 24% in comparison to AI which was 100% accurate. This is a 316% increase in accuracy.

# Analysis

For star B, the average accuracy was 54% in comparison to AI which was 100% accurate. This is an 85% increase in accuracy.

For star C, the average accuracy was 43% in comparison to AI which was 100% accurate. This is a 133% increase in accuracy.

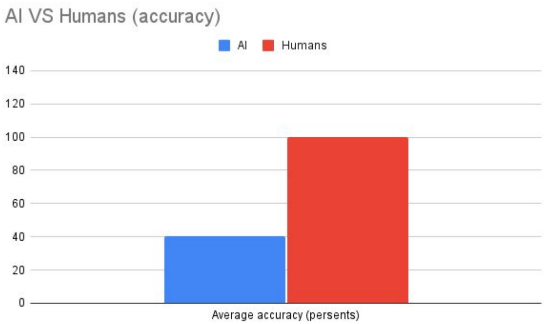
For star D, the average accuracy was 37% in comparison to AI which was 100% accurate. This is a 170% increase in accuracy.

For star E, the average accuracy was 44% in comparison to AI which was 100% accurate. This is a 127% increase in accuracy.

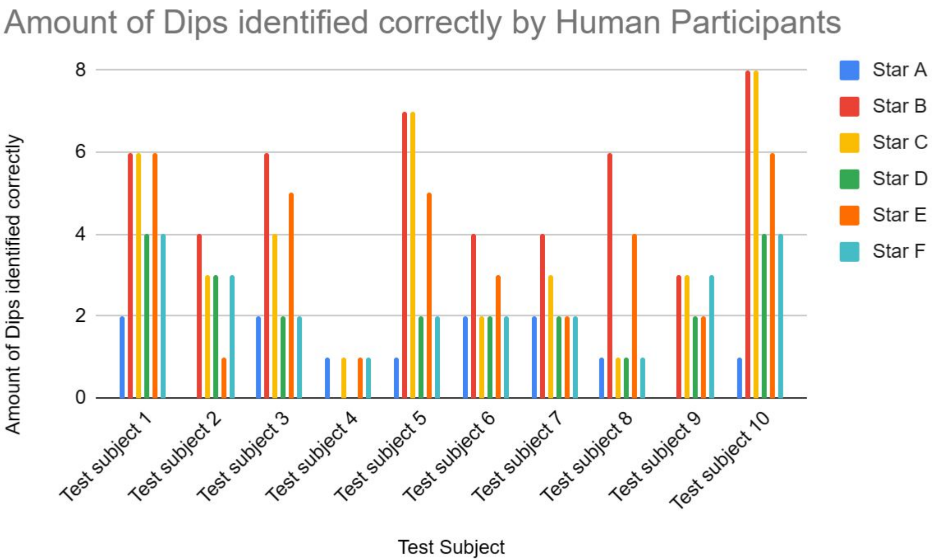
# Analysis

For star F, average accuracy was 40% in comparison to AI which was 100% accurate. This is a 150% increase in accuracy.

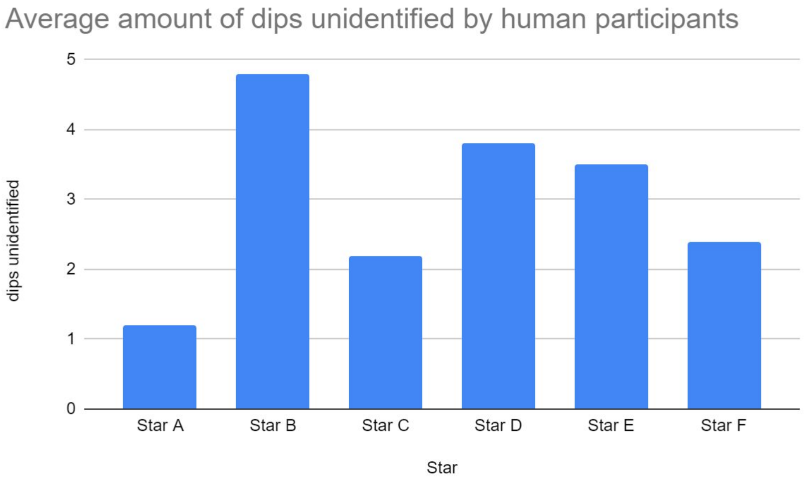
From this information it can be concluded that AI is countless times better than human participants. Below are some more graphs that represent this information:



# Analysis

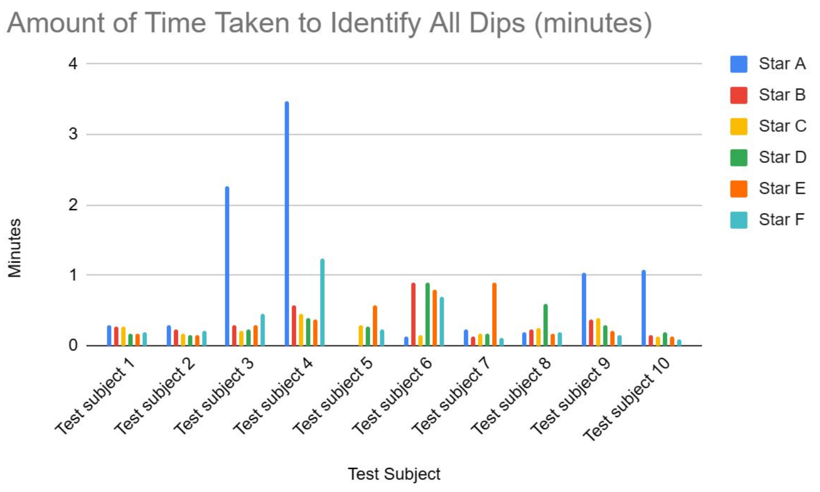


Amount of Dips identified correctly by Human Participants
Amount of Dips identified correctly
0
2
4
6
8
Test subject 1
Test subject 2
Test subject 3
Test subject 4
Test subject 5
Test subject 6
Test subject 7
Test subject 8
Test subject 9
Test subject 10
Test Subject
Star A
Star B
Star C
Star D
Star E
Star F

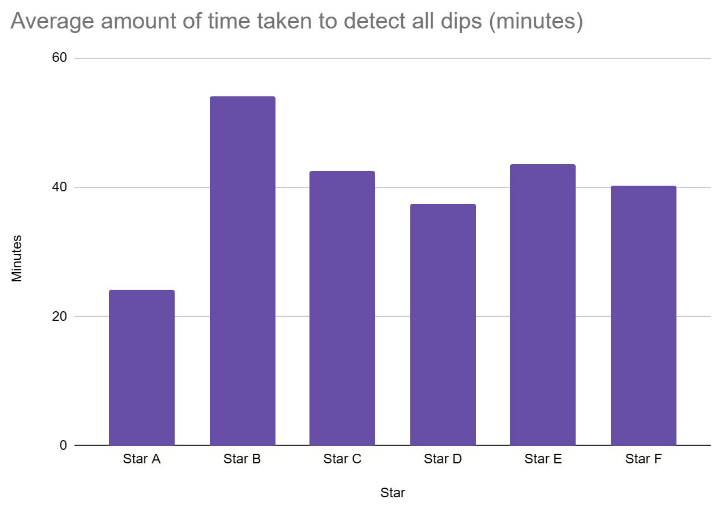



Average amount of dips unidentified by human participants
dips unidentified
0
1
2
3
4
5
Star A
Star B
Star C
Star D
Star E
Star F
Star

# Analysis

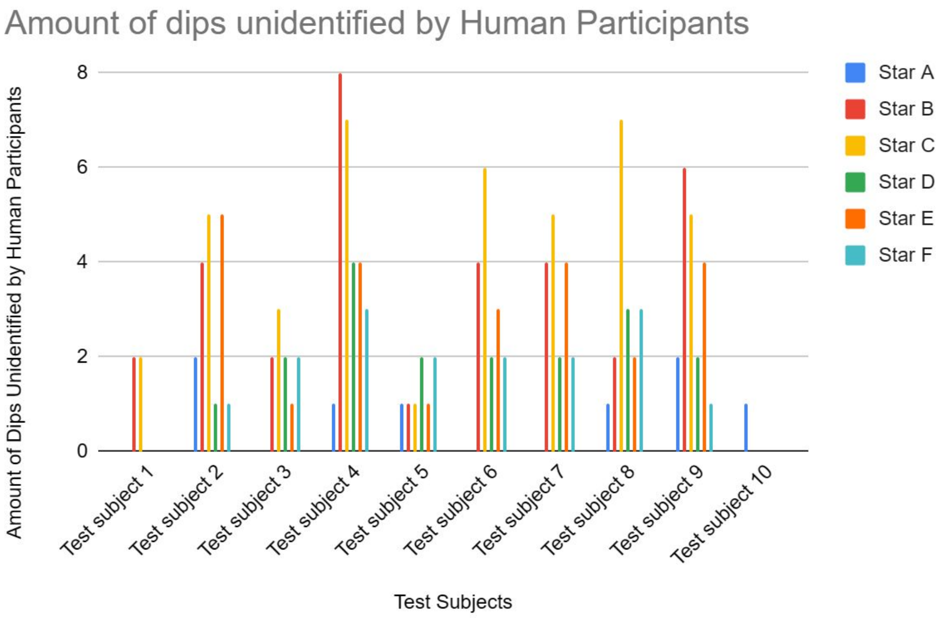


Amount of Time Taken to Identify All Dips (minutes)
Minutes
Test Subject
Test subject 1
Test subject 2
Test subject 3
Test subject 4
Test subject 5
Test subject 6
Test subject 7
Test subject 8
Test subject 9
Test subject 10
0
1
2
3
4
Star A
Star B
Star C
Star D
Star E
Star F

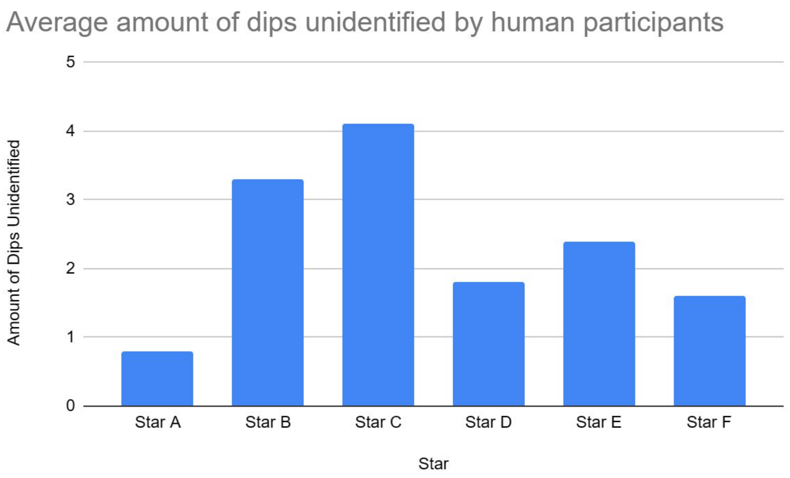



Average amount of time taken to detect all dips (minutes)
Minutes
Star
0
20
40
60
Star A
Star B
Star C
Star D
Star E
Star F

# Analysis

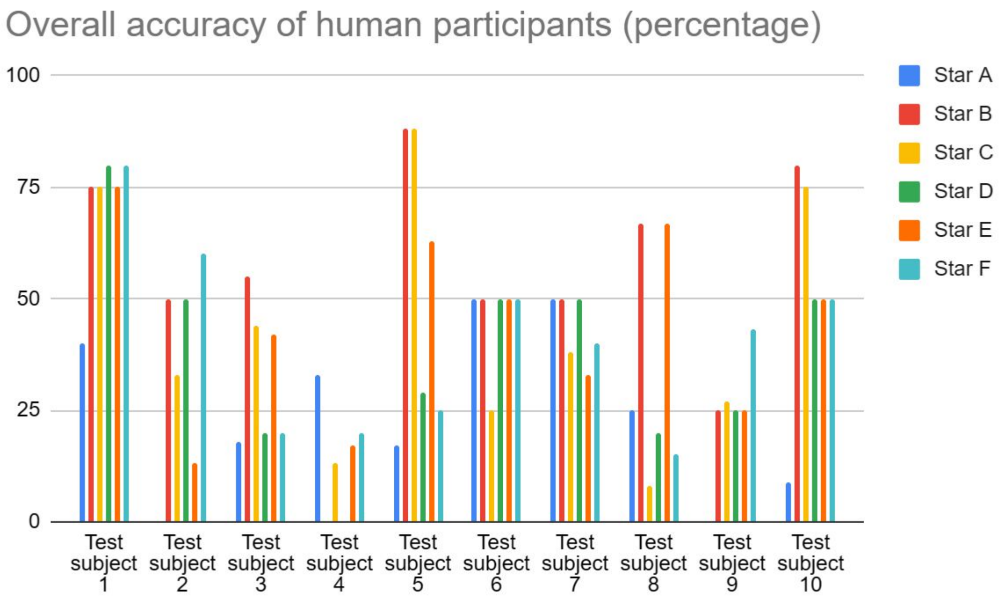


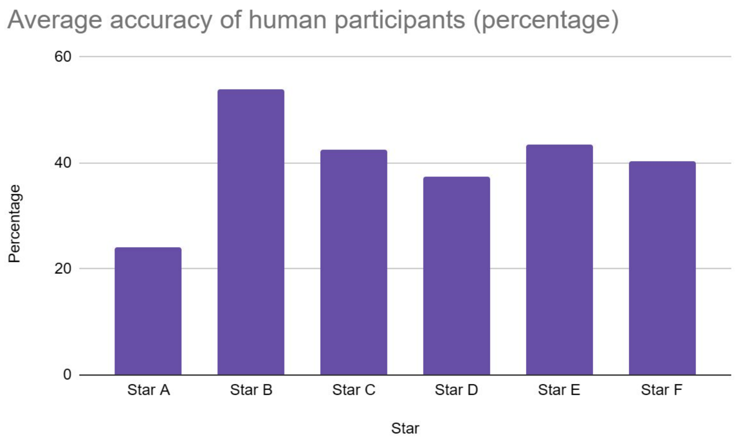
Amount of dips unidentified by Human Participants
Amount of Dips Unidentified by Human Participants
0
2
4
6
8
Test subject 1
Test subject 2
Test subject 3
Test subject 4
Test subject 5
Test subject 6
Test subject 7
Test subject 8
Test subject 9
Test subject 10
Test Subjects
Star A
Star B
Star C
Star D
Star E
Star F

Average amount of dips unidentified by human participants
Amount of Dips Unidentified
0
1
2
3
4
5
Star A
Star B
Star C
Star D
Star E
Star F
Star

# Analysis


Overall accuracy of human participants (percentage)
100
75
50
25
0
Test subject 1
Test subject 2
Test subject 3
Test subject 4
Test subject 5
Test subject 6
Test subject 7
Test subject 8
Test subject 9
Test subject 10
Star A
Star B
Star C
Star D
Star E
Star F


Average accuracy of human participants (percentage)
60
40
20
0
Percentage
Star A
Star B
Star C
Star D
Star E
Star F
Star

# Conclusion

# Conclusion

In conclusion, the original hypothesis that AI can make exoplanet discovery more efficient and accurate is strongly supported by both the experimental and literature review sections.

Traditional methods involving statistical analysis, such as periodicity and Fourier analysis, were found to be time-consuming and extremely complex. However, using deep learning and machine learning models has proven to be much more effective at analysing star brightness data. Some of Artificial Intelligence's abilities include recognition, efficiency, and the ability to analyse a large amount of data.

# Conclusion

Through the experimental portion of this project, it was found that human participants faced several challenges. Many stated that they found "star A" to be the most difficult, as it was the first graph they had to analyse. Like machine learning, human accuracy consistently improved with more data, but it took a considerable amount of time to explain how to analyse the data. Another challenge was the nature of human analysis, where opinions played a significant role. For instance, a dip in a star's brightness could be attributed to various factors, not just an exoplanet. Humans also struggled with the time difference between AI and humans, taking an average of 40 seconds per graph compared to the practically instantaneous task completion by the AI algorithm

# Conclusion

Quantitative data further supports the superiority of AI over humans. The AI algorithm achieved an average accuracy of 100%, while human participants had an accuracy of 42.6%, taking 44% more time to detect all dips. These numbers underscore the previously mentioned points.

The practical applications of this project extend to various fields. Firstly, it demonstrates how AI significantly reduces time and effort by automating the task of identifying dips in scatterplots accurately. This project also illustrates how AI can help detect dips in scatterplots, showcasing its potential in large-scale surveys like TESS. By using AI, the efficiency of detecting exoplanets, and potentially life outside Earth, can be greatly improved.

# Conclusion

In general, this project highlights the potential of AI in optimising resources and dealing with more complex datasets. It also indicates technological advancements, showcasing the potential applications of AI in the future.

While this project is relatively simple, accurately detecting exoplanets with AI through transit photometry data requires further development for a more accurate model.

# Applications

# Applications

This project has countless applications. Firstly, the most obvious application is the fact that we will be able to discover exoplanets. Discovering exoplanets will help humanity in many ways. One being the most prominent is finding life outside of Earth. Discovering extraterrestrial life has many benefits. For example, discovering life will drive scientific curiosity and push us to continue researching and studying space. Finding life outside of Earth will also help to gain understanding about the origin of the human race because it will help answer questions like if life arose independently or if it has a connection with us.

Discovering life outside of Earth will help to expand humanity's perspective of the world and help make advancements in technologies, because life that is not found on Earth may have made advancements that we do not know of yet.  This will help advance knowledge in fields such as chemistry, biology and physics.

# Applications

Another huge benefit of looking for and finding life outside of Earth is important for figuring out the future of Humanity. Will we have to colonise another planet? What will be the future of Earth? Are questions that humans have been asking, especially now that climate change has been a big issue. Finding life can help finding a solution to this problem.

This project emphasises how AI can be used to make various activities take less time and effort as well as improve accuracy of certain tasks. This information is proven in the data that proves how good AI does what it is told to do.

# Applications

There are many ways AI could and is improving the lives of people. Firstly, AI's use in the healthcare industry can improve the accuracy of diagnosing diseases by analysing medical images like x-rays and MRI's.

AI can also revolutionise communication by translating languages. This breaks the communication barrier allowing people to communicate from all over the world.

Artificial intelligence can contribute to reducing the negative impact of humans on planet Earth. This is because it is able to identify where and how humans are able to reduce their carbon footprint. AI can also help with energy optimisation by adjusting heating, cooling, lighting, etc. based on occupancy and other factors.

# Applications

Throughout the project, it was evident that Artificial Intelligence excels at identifying patterns. This prompts new questions, such as:

1. Given that light takes many years to travel from distant galaxies to Earth, essentially allowing us to look back in time, can we determine if our galaxy once had a quasar using data from other galaxies? Can AI assist in this, and if so, how?
2. What are the other potential applications of AI in the field of astronomy?

# Sources of Error

# Sources of Error

- Previous experience of the human participant may have impacted how accurate and how much time they take to analyse a graph. For example, someone working in the field of data analysis may have more knowledge about scatter plots than someone who works in performing arts.

- Though the information given to the participant before testing was the same, a participant may have interpreted it in a different way. For example, one participant thought to identify the points that are darker on colour rather than being lower on the Y-axis.

Because all sources of error are caused by human error, this information actually supports Artificial intelligence and continues to prove why AI should be used.

# Citations

# Citations

NASA. (2021, April 2). *Overview*. NASA. https://exoplanets.nasa.gov/what-is-an-exoplanet/overview/

NASA. (2021, April 2). *Overview*. NASA. https://exoplanets.nasa.gov/what-is-an-exoplanet/overview/

NASA. (2021, March 22). *Transiting Exoplanet Survey Satellite (tess) - exoplanet exploration: Planets beyond our solar system*. NASA. https://exoplanets.nasa.gov/tess/

*Exoplanets, Worlds orbiting other stars*. The Planetary Society. (n.d.). https://www.planetary.org/worlds/exoplanets#:~:text=There%20are%20lots%20of%20reasons,the%20course%20of%20human%20history.

# Citations

*Exoplanet and candidate statistics*. Exoplanet and Candidate Statitics. (n.d.). https://exoplanetarchive.ipac.caltech.edu/docs/counts_detail.html

NASA. (2019, June 20). 5 ways to find a planet. NASA. https://exoplanets.nasa.gov/alien-worlds/ways-to-find-a-planet/

*What is Periodicity analysis?* (n.d.). Www.tutorialspoint.com. Retrieved December 14, 2023, from https://www.tutorialspoint.com/what-is-periodicity-analysis

Das, T. (2023, July 8). *What is Artificial Intelligence (AI)?* Tech4Fresher. https://tech4fresher.com/what-is-artificial-intelligence-ai/