Group 15 CSIRO Presentation

Afeef Ahmed
Alexandra Ellison
Matthew Oxley
Max Grant
Tajkinur Rahman

PULSAR PIONEERS

LOOKING INTO THE FUTURE BY LOOKING INTO THE PAST

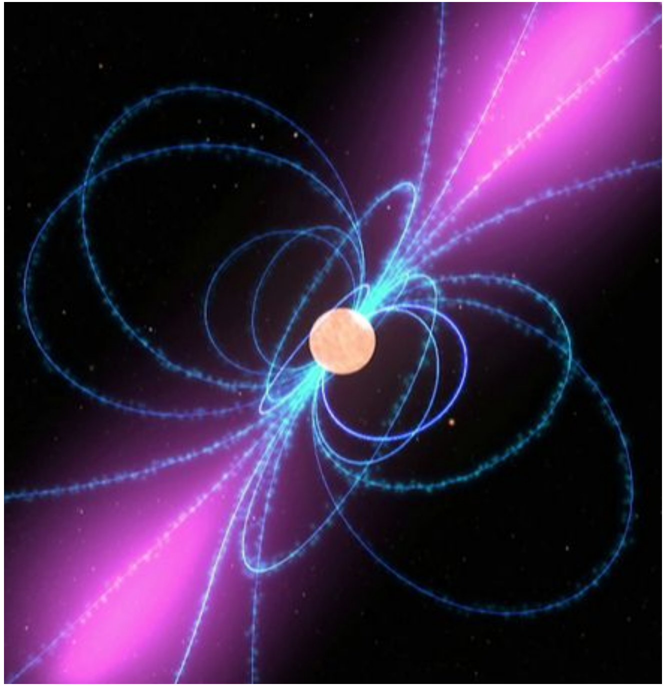
The Importance of Randomness

- Random number generators are important in society
- We want to find an unbiased random number generator that can be accessed by multiple parties
- Random numbers are currently generated with computers, however can be interfered with
- Instead, we want to look at naturally occurring random numbers, that cannot be replicated

Introducing Pulsars

- Pulsars are rapidly spinning stars that emit light beams
- These light pulses are thought to be of random brightness
- Therefore, the measure of brightness might provide us with a random number sequence

- The CSIRO asked us to investigate the brightness data of 6 specific pulsars
- Our aim is to conclude whether the pulses from these pulsars are truly random



Our Task

- Some adjacent pulses are correlated over time, and are evidently not random
- We want to remove these correlated pulses and test for randomness
- Do this by extracting a subset of pulses from the data, as per the calculated autocorrelation function
- Then, test this subset for randomness, eliminating the observed correlation or ‘pattern’

The Data

All the six pulsars datasets undergoes analysis

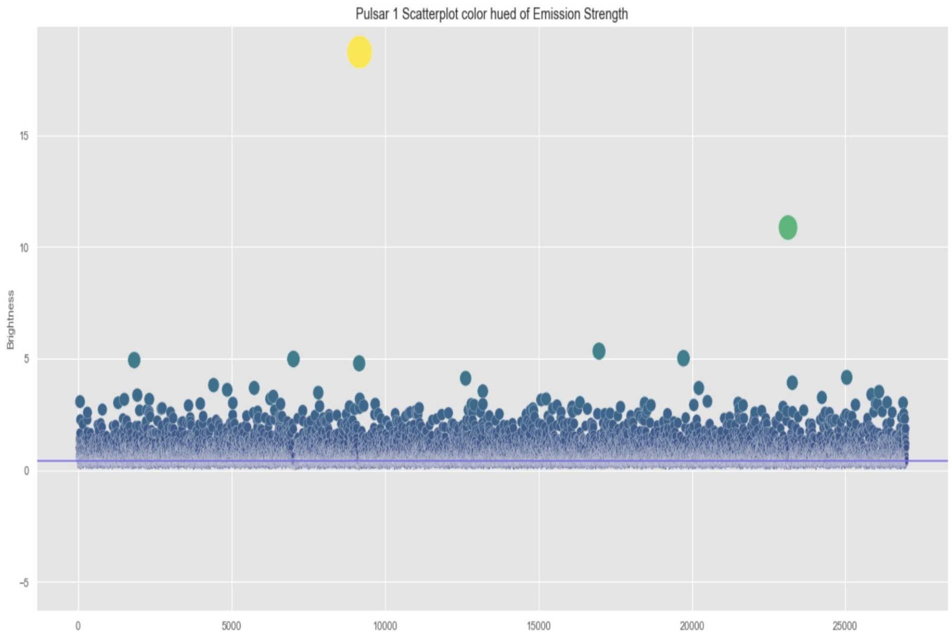
1. J0437-4715
2. J0953+0755
3. J0835-4510
4. J1243-6423
5. J1456-6843
6. J1644-4559

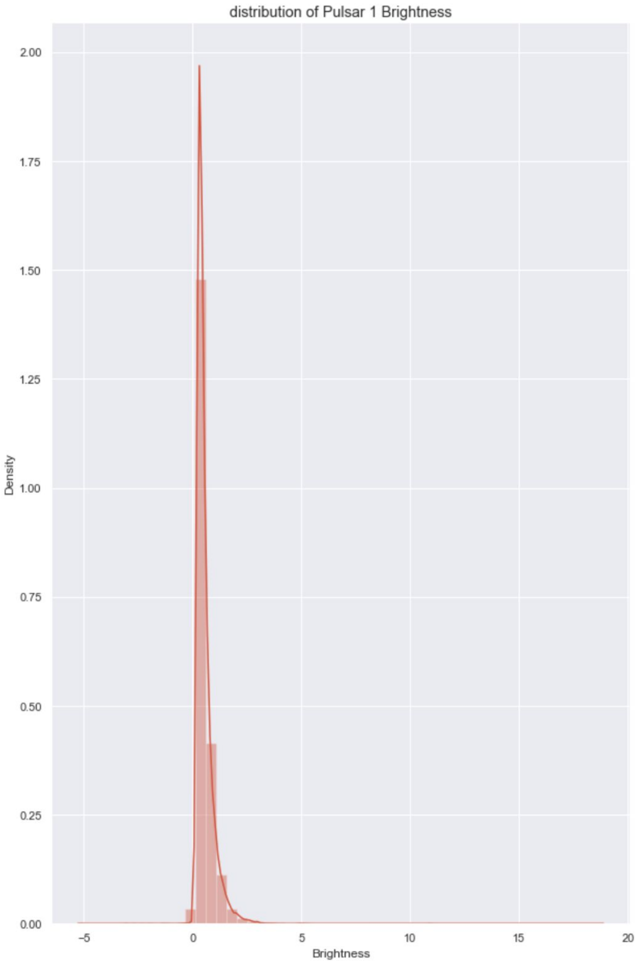
We first start off with examining our dataset and determine if data cleaning and uniforming process is required. This is done to fix any data that has incorrect, inaccurate, incomplete to the data set.

We then carry out various plots to get a visual understanding of the data

Analysis of the scatterplot

Looking at the the scatterplot of the pulsar 1 we can say that the data points are randomly scattered and there is no pattern seen. We can observe that there are two data points that have an unusual high value compared to the rest. As we measure the median, these outliers are unlikely to impact the results of the test.





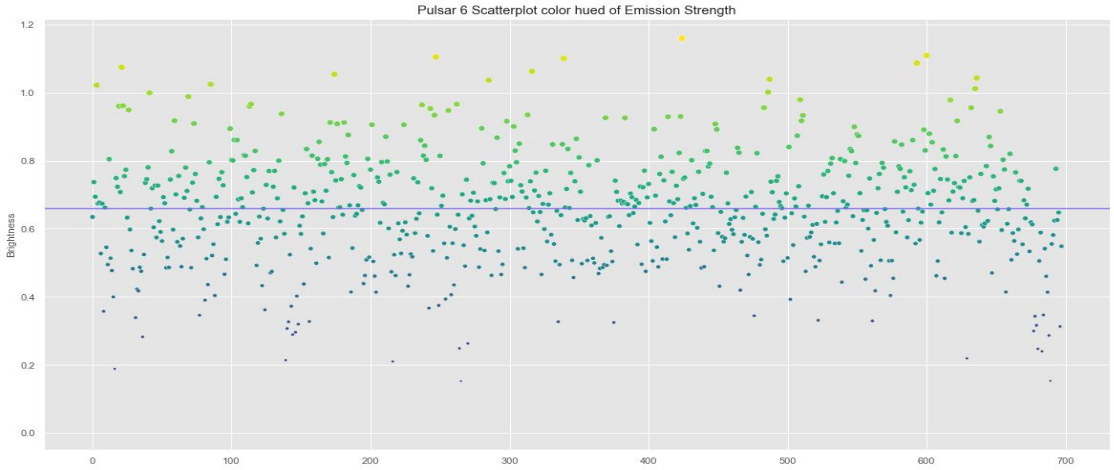
Distribution of the Plot

We check the distribution of pulsar where it tells us that there is no normality and we observe a positive right skewed distribution and most of values on brightness scale fall close to 0 with the largest brightness value around 3.

Analysis of the scatterplot

For this scatterplot, we observe the data points are randomly scattered and there is no observable pattern here as well.

The data points are divided into 1’s and 0’s for binary classification where the data points below the median line are 0 and above the median line is 1



Binary Column

For our randomness tests, we need a binary sequence. We added a ‘Binary’ column which was established using the brightness median as a centre point. If a Brightness value was greater than the median it was assigned 1 if less than it was 0. This is the sequence we used for the randomness tests.

	Pulse Number	Brightness	Uncertainty	Binary
0	1	0.598393	0.056431	1
1	2	0.590859	0.055182	1
2	3	0.449643	0.063632	1
3	4	0.682860	0.056269	1
4	5	0.490026	0.046830	1
5	6	0.586071	0.052649	1
6	7	0.150353	0.056483	0

Fig: Binary column with “Brightness” median being 0.4238

Autocorrelation

Autocorrelation represents the similarity between a given time series and a lagged version of itself over successive time intervals. Basically it measures the relationship between a variable's current value and its past values. If we check the autocorrelation of this dataset's brightness value, we can see that the interval is not statistically significant after 5. Therefore, a selection of every 5th is made.

```
plt.style.use("seaborn")
plt.rcParams["figure.figsize"] = (18, 9)

fig, ax = plt.subplots(2,1)

plot_acf(pulsar['Brightness'], ax=ax[0])
plot_pacf(pulsar['Brightness'], ax=ax[1], method="ols")
```

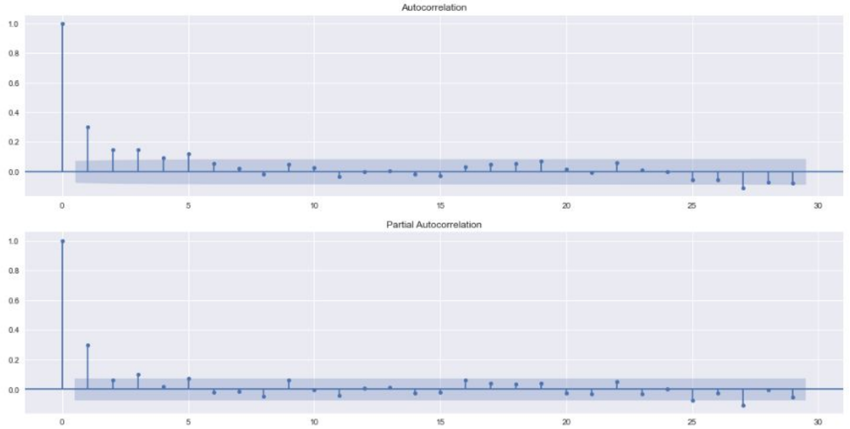


Fig: Pulsar 3 ACF

Machine Learning

- **Binary Classification:** Classification problems with two class labels are referred to as binary classification. We used binary classification to predict the binary value of the data. The model predicted the values with more than optimal accuracy, recall, precision and F1 score which means the model is performing well but this did not help us further analyse the data.

In [39]:
```
# accuracy, recall, precision and F1

acc = (TP + TN) / (TP + FP + TN + FN)
recall = (TP)/(FN+TN)
precision = (TP)/(TP+FP)
f1 = (2)/((1/recall)+ (1/precision))

print("Accuracy: ", acc)
print("Recall: ", recall)
print("Precision: ",precision)
print("F1 Score: ", f1)
```

```
Accuracy:  0.997037037037037
Recall:  0.9834680382072006
Precision:  0.9996265870052278
F1 Score:  0.9914814814814814
```

Machine Learning

- **Bidirectional LSTM Model:** Bidirectional LSTM is the process of making any neural network to have the sequence information in both directions. This model helped us with the initial assessment of the data to see if it follows a certain trend/sequence. This model is often used to test pseudo-random numbers and it works significantly well. In this case, we used the “Brightness” and the “Uncertainty” values to train the model and the results showed us that no matter how many previous values we feed the value to predict the next value, it cannot find a sequence to correctly predict the correct value.

```
In [51]: # evaluating the model
         model.evaluate(X_test, y_test)

         169/169 [==============================] - 2s 11ms/step - loss: 0.0019 - accuracy: 0.0000e+00

Out[51]: [0.0018631333950906992, 0.0]
```

The Randomness Tests

16 different tests in NIST test suite

- **Frequency Test (Monobit)** → equal proportions of 1’s and 0’s
- **Frequency Test within a Block** → equal proportion of 1’s and 0’s in block
- **Run Test** → total number of runs of k identical bits
- **Longest Run of Ones in a Block**
- **Binary Matrix Rank Test** → tests for linear dependence of substrings of the sequence

1 0 1 0 0 1 1 1 0 0 1 1 0 0 10

The Results

Pulsar 1

Every Observation	
Frequency Test (Monobit)	Pass
Frequency Test within a Block	Fail
Run Test	Fail
Longest Run of Ones in a Block	Pass
Binary Matrix Rank Test	N/A
Total Passes	**3**

Every 5th Observation	
Frequency Test (Monobit)	Pass
Frequency Test within a Block	Pass
Run Test	Pass
Longest Run of Ones in a Block	Pass
Binary Matrix Rank Test	Pass
Total Passes	**5**

α = 0.01

The Results

Pulsar 2

Every Observation	
Frequency Test (Monobit)	Pass
Frequency Test within a Block	Fail
Run Test	Fail
Longest Run of Ones in a Block	Fail
Binary Matrix Rank Test	Pass
Total Passes	**2**

Every 5th Observation	
Frequency Test (Monobit)	Pass
Frequency Test within a Block	Pass
Run Test	Pass
Longest Run of Ones in a Block	Pass
Binary Matrix Rank Test	Pass
Total Passes	**5**

α = 0.01

The Results

Pulsar 3

Every Observation	
Frequency Test (Monobit)	Pass
Frequency Test within a Block	Pass
Run Test	Pass
Longest Run of Ones in a Block	Pass
Binary Matrix Rank Test	Pass
Total Passes	**5**

Every 5th Observation	
Frequency Test (Monobit)	Pass
Frequency Test within a Block	Pass
Run Test	Pass
Longest Run of Ones in a Block	Pass
Binary Matrix Rank Test	N/A
Total Passes	**4**

α = 0.01

The Results

Pulsar 4

Every Observation	
Frequency Test (Monobit)	Pass
Frequency Test within a Block	Fail
Run Test	Fail
Longest Run of Ones in a Block	Fail
Binary Matrix Rank Test	Pass
Total Passes	**2**

Every <u>10th</u> Observation	
Frequency Test (Monobit)	Pass
Frequency Test within a Block	Pass
Run Test	Fail
Longest Run of Ones in a Block	Pass
Binary Matrix Rank Test	N/A
Total Passes	**3**

α = 0.01

The Results

Pulsar 5

Every Observation	
Frequency Test (Monobit)	Pass
Frequency Test within a Block	Pass
Run Test	Fail
Longest Run of Ones in a Block	Pass
Binary Matrix Rank Test	Pass
Total Passes	**4**

Every 5th Observation	
Frequency Test (Monobit)	Pass
Frequency Test within a Block	Pass
Run Test	Pass
Longest Run of Ones in a Block	Pass
Binary Matrix Rank Test	N/A
Total Passes	**4**

α = 0.01

The Results

Pulsar 6

Every Observation	
Frequency Test (Monobit)	Pass
Frequency Test within a Block	Pass
Run Test	Fail
Longest Run of Ones in a Block	Fail
Binary Matrix Rank Test	N/A
Total Passes	**2**

Every 5th Observation	
Frequency Test (Monobit)	Pass
Frequency Test within a Block	Pass
Run Test	Pass
Longest Run of Ones in a Block	Pass
Binary Matrix Rank Test	N/A
Total Passes	**4**

α = 0.01

Which Pulsars are random?

- From the previous slides we can confidently rule out using the 5 tests that we have interpreted as the most important. We can start to begin to either rule out randomness in a pulsar or not.

Pulsar observations ALL EMISSIONS

	Test Name	Outcome	Outcome2	Outcome3	Outcome4	Outcome5	Outcome6
0	Frequency Test (Monobit)	Absolute	Random	Random	Random	Random	Absolute
1	Frequency Test within a Block	Non-Random	Non-Random	Random	Non-Random	Random	Random
2	Run Test	Non-Random	Non-Random	Random	Non-Random	Non-Random	Non-Random
3	Longest Run of Ones in a Block	Random	Non-Random	Random	Non-Random	Random	Non-Random
4	Binary Matrix Rank Test	Random	Random	Random	Random	Random	Error
5	Discrete Fourier Transform (Spectral) Test	Random	Non-Random	Random	Random	Random	Random
6	Non-Overlapping Template Matching Test	Non-Random	Non-Random	Random	Random	Random	Random
7	Overlapping Template Matching Test	Non-Random	Non-Random	Random	Random	Random	Error
8	Maurer's Universal Statistical Test	Error	Error	Error	Error	Error	Error
9	Linear Complexity Test	Random	Random	Random	Random	Random	Error
10	Serial Test A	Non-Random	Non-Random	Random	Non-Random	Random	Random
11	Serial Test B	Non-Random	Random	Random	Non-Random	Random	Random
12	Approximate Entropy Test	Non-Random	Non-Random	Random	Non-Random	Random	Absolute
13	Cumulative Sums (Forward) Test	Random	Random	Random	Non-Random	Random	Random
14	Cumulative Sums (Reverse) Test	Random	Random	Random	Non-Random	Random	Random
15	Random Excursions Test State -4	Random	Random	Random	Random	Non-Random	Random
16	Random Excursions Test State -3	Random	Non-Random	Random	Random	Random	Random
17	Random Excursions Test State -2	Random	Non-Random	Random	Random	Random	Random
18	Random Excursions Test State -1	Random	Random	Random	Random	Random	Random
19	Random Excursions Test State +1	Random	Random	Random	Random	Random	Random
20	Random Excursions Test State +2	Random	Random	Random	Random	Random	Random
21	Random Excursions Test State +3	Random	Random	Random	Random	Random	Random
22	Random Excursions Test State +4	Random	Random	Random	Random	Random	Non-Random
23	Random Exursions Variant Test State -9	Random	Random	Random	Random	Random	Random
24	Random Exursions Variant Test State -8	Random	Random	Random	Random	Random	Random
25	Random Exursions Variant Test State -7	Random	Random	Random	Random	Random	Random
26	Random Exursions Variant Test State -6	Random	Random	Random	Random	Random	Random
27	Random Exursions Variant Test State -5	Random	Random	Random	Random	Random	Random
28	Random Exursions Variant Test State -4	Random	Random	Random	Random	Random	Random
29	Random Exursions Variant Test State -3	Random	Random	Random	Random	Random	Random
30	Random Exursions Variant Test State -2	Random	Random	Random	Random	Random	Random
31	Random Exursions Variant Test State -1	Random	Random	Random	Random	Random	Random
32	Random Exursions Variant Test State +1	Random	Random	Random	Random	Random	Random
33	Random Exursions Variant Test State +2	Random	Random	Random	NaN	Random	Random
34	Random Exursions Variant Test State +3	Random	Random	Random	NaN	Random	Random
35	Random Exursions Variant Test State +4	Random	Random	Random	NaN	Absolute	Random
36	Random Exursions Variant Test State +5	Random	Random	Random	NaN	Random	Random
37	Random Exursions Variant Test State +6	Random	Random	Random	NaN	Random	Random
38	Random Exursions Variant Test State +7	Random	Random	Random	NaN	Random	Random
39	Random Exursions Variant Test State +8	Random	Random	Random	NaN	Random	Random
40	Random Exursions Variant Test State +9	Random	Random	Random	NaN	Random	Random

RandTestAll
True
True
True
True
True
True

From both RandTest and NIST Suite we can say that.

Pulsars 3 & 5 are random - requires some more investigation into the errors and more. 5 could potentially be not random due to failure of the runs test (A fundamental test of randomness.)

EVERY 5TH & 10th EMISSION

	Test Name	Outcome	Outcome2	Outcome3	Outcome4	Outcome5	Outcome6
0	Frequency Test (Monobit)	Random	Random	Random	Random	Random	Random
1	Frequency Test within a Block	Random	Random	Random	Non-Random	Random	Random
2	Run Test	Random	Random	Random	Non-Random	Random	Random
3	Longest Run of Ones in a Block	Random	Random	Random	Random	Random	Random
4	Binary Matrix Rank Test	Random	Random	Error	Error	Error	Error
5	Discrete Fourier Transform (Spectral) Test	Random	Random	Non-Random	Random	Random	Random
6	Non-Overlapping Template Matching Test	Random	Random	Random	Absolute	Absolute	Random
7	Overlapping Template Matching Test	Random	Random	Error	Error	Error	Error
8	Maurer's Universal Statistical Test	Error	Error	Error	Error	Error	Error
9	Linear Complexity Test	Random	Random	Error	Error	Error	Error
10	Serial Test A	Random	Random	Random	Non-Random	Random	Random
11	Serial Test B	Random	Random	Random	Non-Random	Random	Random
12	Approximate Entropy Test	Non-Random	Non-Random	Absolute	Absolute	Absolute	Absolute
13	Cumulative Sums (Forward) Test	Random	Random	Random	Non-Random	Random	Random
14	Cumulative Sums (Reverse) Test	Random	Random	Random	Non-Random	Random	Random
15	Random Excursions Test State -4	Random	Random	Random	Random	Random	Random
16	Random Excursions Test State -3	Random	Random	Random	Random	Random	Random
17	Random Excursions Test State -2	Random	Random	Random	Random	Random	Random
18	Random Excursions Test State -1	Random	Random	Random	Random	Random	Random
19	Random Excursions Test State +1	Random	Non-Random	Random	Non-Random	Random	Random
20	Random Excursions Test State +2	Non-Random	Non-Random	Random	Non-Random	Random	Random
21	Random Excursions Test State +3	Non-Random	Random	Random	Non-Random	Random	Random
22	Random Excursions Test State +4	Random	Random	Random	Non-Random	Random	Random
23	Random Exursions Variant Test State +1	Random	Random	Random	Random	Random	Random
24	Random Exursions Variant Test State +2	Random	Random	Random	Non-Random	Random	Random
25	Random Exursions Variant Test State +3	Random	Random	Random	Non-Random	Random	Random
26	Random Exursions Variant Test State +4	Non-Random	Random	Random	Non-Random	Random	Random
27	Random Exursions Variant Test State +5	Non-Random	Random	Random	Random	Random	Random
28	Random Exursions Variant Test State +6	Non-Random	Random	Random	Random	Random	Random
29	Random Exursions Variant Test State +7	Non-Random	Random	Random	Random	Random	Random
30	Random Exursions Variant Test State +8	Random	Random	Random	Random	Random	Random
31	Random Exursions Variant Test State +9	Random	Random	Random	Random	Random	Random

Outcome
Random
Random
Non-Random
Random
Error
Random
Absolute
Error
Error
Error
Non-Random
Non-Random
Absolute
Random
Random
Random
Random
Random
Random
Random
Random
Random
Random
Random
Random
Random
Absolute
Random
Random
Random
Random
Random
Random
Random
Random

From both RandTest and NIST Suite we can say that.

Pulsars 1, 2, 3, 5 & 6 are likely to be random. Required some more investigation into the errors.

RandTest5ths	RandTestElse
True	Null
True	Null
True	Null
True	True
True	Null
True	Null

Summary Table

Designation	Assignment	All Emissions	Every 5th Emission	Every 10th Emission
J0437-4715	**Pulsar 1**	No	Yes	N/A
J0953+0755	**Pulsar 2**	No	Yes	N/A
J0835-4510	**Pulsar 3**	Yes	Yes	N/A
J1243-6423	**Pulsar 4**	No	No	Yes
J1456-6843	**Pulsar 5**	Yes	Yes	N/A
J1644-4559	**Pulsar 6**	No	Yes	N/A

Yes = Random | No = Non-Random

What didn’t work in the project?

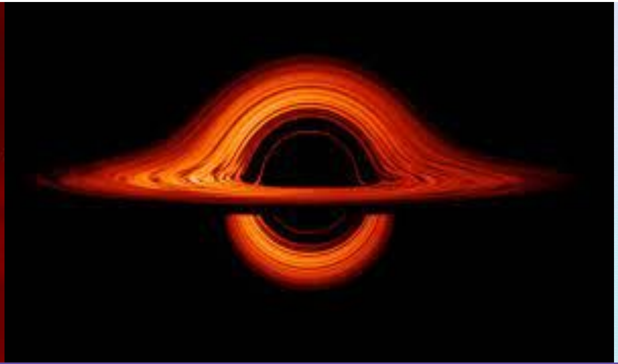
- Kalman Filter didn’t work couldn’t get the coding to satisfy
- Couldn’t get the binned statistics and rolling means and medians to find rewarding data.
- Could not derive a dataset that could satisfy all NIST Tests to produce no errors.

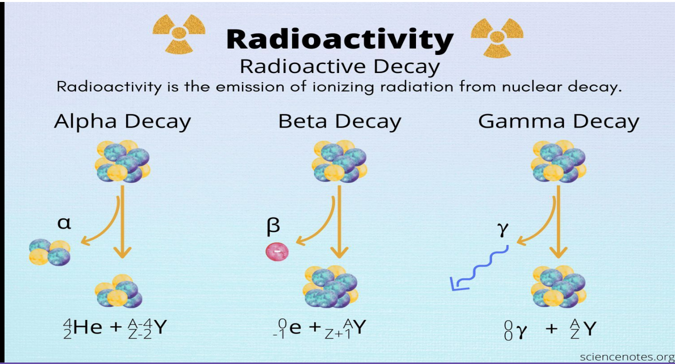
Further Recommendation for this study

- More Pulsars and/or more observations
- Analysis of different testing methods such as TestU01 and ent utility.
- Further investigation of uncertainty.

Further Recommendation for finding Randomness

- Areas where we could find more randomness are abundant.
- Sunspots and coronal mass ejections. (based on emission strength and appearance not location)
- Black Hole Accretion disks - emissions of colliding objects
- Radioactive Decay based on the specific atom not rate.





CONCLUSION

Through this project we believe that we are confident in the areas we have undertaken:

- Our Data analysis
- Attempts of various methods to reward more data
- Testing and interpretations
- Results

THANK YOU!

Afeef Ahmed

Alexandra Ellison

Matthew Oxley

Max Grant

Tajkinur Rahman

PULSAR PIONEERS

LOOKING INTO THE FUTURE BY
LOOKING INTO THE PAST