The Riemann Hypothesis: A Fresh and Experimental Exploration

Abstract. This research proposes a new approach to the Riemann Hypothesis, focusing on the interplay between prime gaps and the non-trivial zeros of the Riemann Zeta function. Utilizing various statistical models and experimental analysis techniques, three important insights are uncovered: 1) Granger causality tests reveal a predictive relationship in which past non-trivial zeros may predict future prime gaps; 2) Complex, nonlinear interactions between prime gaps and non-trivial zeros are identified, challenging simple linear correlations; and 3) Causal network analysis reveals intricate feedback-loop relationships. These findings contribute to a better understanding of prime number distribution and the Zeta function, opening up novel possibilities for further mathematical research.The study aims to motivate mathematicians towards a proof or disproof of the Riemann Hypothesis.

Keywords: Riemann Hypothesis, prime gaps, non-trivial zeros of the Zeta function, statistical models.

1. Introduction

The Riemann Hypothesis, proposed in 1859 by German mathematician Bernhard Riemann, is a renowned and enduring problem in mathematics, focusing on prime number distribution. It is a key element in number theory, deeply influencing our understanding of primes. The Riemann Hypothesis concerns the zeros of the Riemann Zeta function, which is closely linked to prime numbers.The Riemann Zeta function, $\zeta(s)$, is a complex function defined for s with Re(s) > 1 by the series $\sum_{n=1}^{\infty}\frac{1}{n^s}$, extending to all complex numbers except s=1 through analytic continuation, and deeply connected to prime numbers through the Euler product formula $\prod_{p\ \text{prime}}\frac{1}{1-p^{-s}}$. The Riemann Zeta function conjectures that all non-trivial zeros of $\zeta(s)$ have a real part of ½. These zeros, complex numbers, are key to grasping how prime numbers are distributed among integers.

The distribution of prime numbers, central in number theory, has wide applications in mathematics and computer science, notably in cryptography. The Riemann Hypothesis is crucial because it suggests profound, extensive insights into prime number distribution beyond current provable knowledge without it. Despite strong numerical support, the Riemann Hypothesis remains unproven. Solving this 150-year-old problem would be a major mathematical breakthrough. The Clay Mathematics Institute offers a $1 million Millennium Prize for proving or disproving it, highlighting its significance and challenge.

This paper presents an innovative exploratory approach to understanding the Riemann Hypothesis, focusing on the relationship between prime gaps and the non-trivial zeros of the Riemann Zeta function. It employs various statistical models and experimental analysis techniques. The study's goal is to motivate mathematicians to discover an ultimate proof of disproof of the Riemann Hypothesis. The difficulties we had in establishing obvious causal links and in adequately modeling the complex interactions examined in this research should serve as inspiration.

Proposed work

Matsushita and Da Silva's study [1]analyzed prime gaps (intervals between consecutive primes) and found they follow a power law, suggesting a non-random pattern. This pattern indicates prime gaps are inversely related to a number's likelihood of being prime. This result aids in understanding the k-tuple conjecture about non-random prime distributions [2], revealing biases in consecutive prime numbers that question their assumed randomness. Although not directly tackling the Riemann Hypothesis, such findings offer a statistical physics view of prime distributions [3–13]. This perspective could indirectly assist Riemann Hypothesis research by providing new insights into prime numbers and their intervals, potentially enriching its analysis with the non-random nature of prime gaps.

This research investigates how the identified power law pattern in prime gaps connects to the Riemann Hypothesis. Instead of relying solely on mathematics, an experimental, exploratory approach is employed.

The link between prime number distribution and the Riemann Hypothesis is through the Riemann Zeta function, $\zeta(s)$. For real parts greater than 1, $\zeta(s)$ is both a sum over integers and a product over primes. This connection is key to understanding prime distribution, as shown in the Prime Number Theorem, which connects the count of primes less than a number x to $x/\log(x)$. Linking the non-random prime gaps [1] to the Riemann Hypothesis, which focuses on the zeros of the Zeta function and prime distribution, is complex. While prime gap analysis is insightful, directly connecting it to the Hypothesis needs a new theoretical framework or math approach that relates these gaps to the Zeta function's zeros. Instead, an exploratory approach for gaining additional insights is used.

Related literature

Recent innovative methods propose new ways to address the Riemann Hypothesis, including proof or counter-example strategies. BjörnTegetmeyer's [14] approach to the Riemann Hypothesis involves analyzing the integral representation of the Zeta function to establish that all non-trivial zeros have a real part of ½. His method focuses on solving integrals linked to the Zeta function's real part using complex mathematical techniques.

Another approach is using quantum field theory to interpret the Riemann Zeta function. Grant N. Remmen [15] identified a scattering amplitude, mirroring the Riemann Zeta function's non-trivial zeros, which describes two massless particles interacting. If this amplitude consistently aligns with a quantum field theory having real-number masses, it could validate the Hypothesis. Remmen'stechnique combines mathematics and physics, introducing novel perspectives and tools for understanding the Zeta function.

These efforts to solve the Riemann Hypothesis showcase a range of strategies, blending classic mathematical analysis with innovative, interdisciplinary techniques. This reflects the dynamic nature of mathematical research and its interaction with scientific areas. However, the Riemann Hypothesis remains unsolved, and these methods are part of the wider, ongoing exploration in the field.

Remmen's [15] research connects the Zeta function's properties to quantum field theory. Therefore, his approach suggests interpreting prime gaps and the non-trivial zeros of the Zeta function, key to the Riemann Hypothesis, in terms of quantum field theory phenomena. Formulating a model linking prime numbers to quantum concepts could offer new insights into prime distribution and the Zeta function's zeros. This interdisciplinary model, bridging statistics in prime gaps with quantum theory interpretations, could be a novel way to address the Riemann Hypothesis.

Exploring how prime gap patterns influence the Zeta function's integral representation, as Tegetmeyer [14] studies, could reveal its zeros' nature. His approach entails proving a correlation between prime gap patterns and the Zeta function's integral behavior, focusing on how these patterns might dictate the non-trivial zeros' real parts being ½. (Statistical correlation measures the strength and direction of a linear relationship between two quantitative variables.) Success hinges on deeply analyzing this relationship, aligning prime gaps with the Zeta function's integral characteristics.His study points us in two paths for further research on the Riemann Zeta function. First, it considers whether prime gap patterns relate to the density and distribution of the Zeta function's non-trivial zeros along its critical line ($\text{Re}(s) = ½$). This might involve statistical analysis or new conjectures about prime gaps and zero distribution. Second, it explores the potential influence of prime gap patterns on the integral form of the Zeta function, especially those parts representing its non-trivial zeros. This would require in-depth analysis of the integral forms and could employ computational simulations to empirically test and visualize theoretical connections.

Paper structure

The structure of this paper is as follows: Section 2 describes the materials and methods employed in the study. Section 3 covers the exploratory analysis that was undertaken. Section 4 presents a discussion of these findings. Section 5 concludes the paper with the final remarks.

2. Materials and methods

Given the previous discussion, the ultimate goal of this exploratory approach is to identify any significant correlations or patterns that could provide insights into the relationship between prime gaps and Riemann Zeta function zeros, thus contributing to a broader understanding of these fundamental mathematical concepts.A dataset of one hundrednon-trivial zeros of the Riemann Zeta function obtained fromhttps://www.lmfdb.org/zeros/zeta/ is utilized and the gaps between the initial one hundred prime numbers are computed.The next step performs a comparison analysis and uses statistical correlation tools to see if there are any relationships between these two sets of data.

Table 1. Comparative descriptive statistics for distributions of prime number gaps and non-trivial zeros of the Zeta function.

Statistic	*Prime gaps*	*Non-trivial zeros of Zeta function*
Count	99	100
Mean	5.44	139.72
Standard deviation	3.38	61.25
Minimum	1	14.13
25th percentile	2	91.57
50th percentile (median)	4	91.57
75th percentile	6	144.56
Maximum	18	236.52
Skewness	1.22	–0.23
Excess kurtosis	1.40	–1.05

The skewness and excess kurtosis describe the distributions' shapes. Skewness near zero implies symmetry. Excess kurtosis, measured against a normal distribution's kurtosis of zero, indicates tail weight: negative for lighter tails, positive for heavier. The prime gaps show positive skewness, suggesting a distribution with tails extending towards larger gaps, and a positive excess kurtosis, indicating heavier tails than a normal distribution. For the non-trivial zeros of the Zeta function, the near-zero skewness denotes symmetry, and the negative excess kurtosis reflects lighter tails compared to a normal distribution.

Figures 1 and 2's histograms clearly show distinct differences in the distributions of prime gaps and non-trivial zeros of the Zeta function. Prime gaps cluster at lower values, with frequency decreasing as gap size increases, typical of prime gap distributions where smaller gaps prevail. In contrast, the non-trivial zeros of the Zeta function are dispersed over a wider range, lacking concentration at lower values, resulting in a more uniformly distributed range.

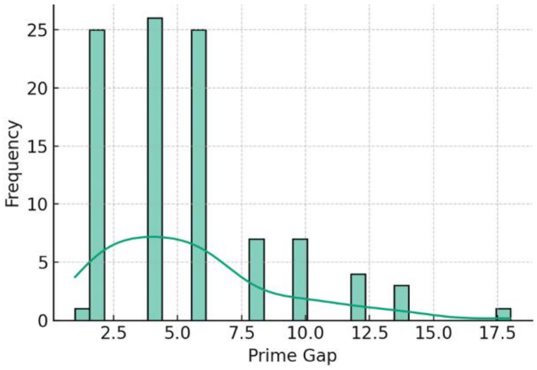

Figure 1. Distribution of prime gaps.

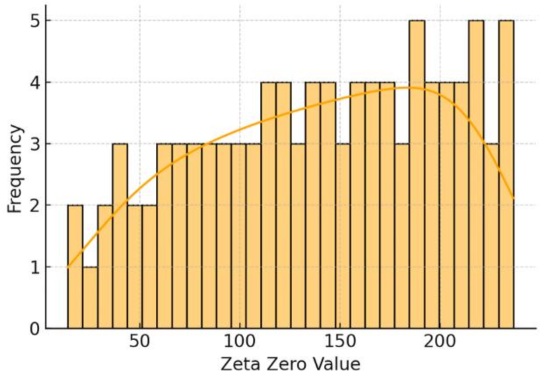

Figure 2. Distribution of the non-trivial zeros of the Zeta function.

The prime gap distribution's positive excess kurtosis aligns with our earlier discovery of a power law in prime gaps [1]. Conversely, the negative excess kurtosis in the non-trivial zeros of the Zeta function suggests the absence of a power law in its distribution.Power law distributions typically look like straight lines on log-log plots because a power law equation $y = ax^k$ becomes linear in logarithmic terms: $\log y = k \log x + \log a$. However, the log-log plot in Figure 3 does not display a straight line. This implies that the non-trivial zeros of the Riemann Zeta function may not adhere to a simple power law. Power laws are known for their heavy tail, meaning larger values are rarer but have a significant impact on the distribution.

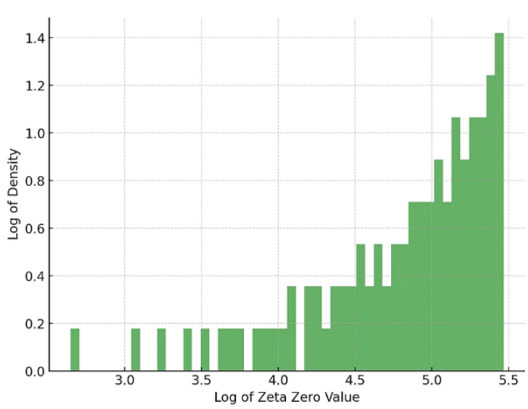

Figure 3. Log-log histogram of non-trivial zeros of the Riemann Zeta function.

Given the unique characteristics of these distributions, a basic correlation analysis might not be effective. Instead, an exploratory experimental analysis to uncover any less obvious relationships or patterns between the two datasets is conducted.

The first step can be looking at the spacing between the zeros of the Zeta function and comparing it to the prime gaps to see if there is any pattern or alignment. Figure 4 depicts a visual comparison of the prime gaps and the spacings between the Zeta function's non-trivial zeros. The scatter plot shows that prime gaps are typically smaller and have a tendency to be smaller, with rare greater gaps. This corresponds to the well-known feature of primes, in which smaller gaps are more prevalent. The Zeta function's zero spacings appear more dispersed and less concentrated than the prime gaps. This implies a more consistent distribution over the value range.

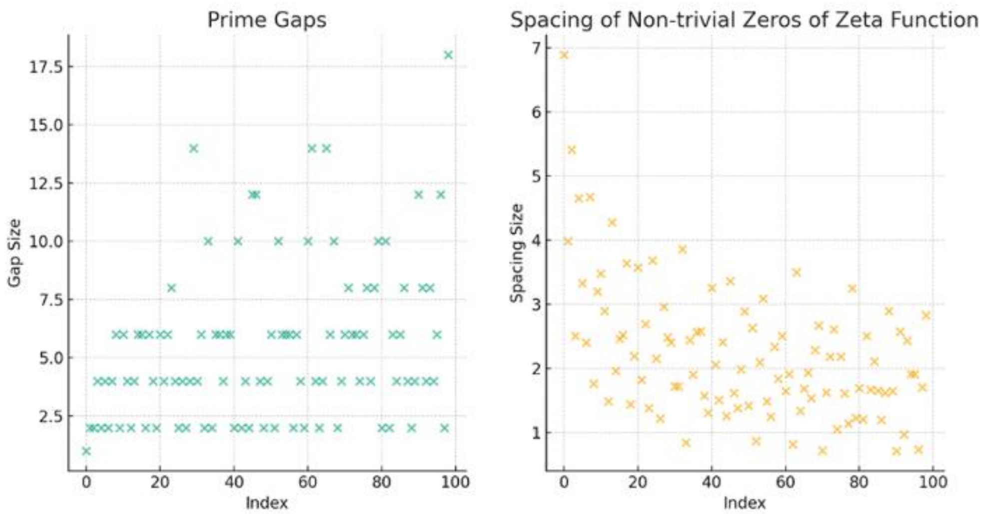

Figure 4. Visual comparison of the prime gaps and the spacings between the non-trivial zeros of the Zeta function.

The correlation coefficient of about −0.05 between the two spacing series suggests a slight inverse relationship. (The correlation coefficient ranges from −1 (perfect negative linear relationship) to 1 (perfect positive linear relationship), with 0 indicating no linear relationship.) However, caution is needed in interpreting this, as the datasets differ fundamentally and direct comparison may not provide significant insights due to their distinct mathematical properties.

The Kolmogorov-Smirnov (KS) test, comparing the cumulative distribution functions (CDFs) of two datasets, reveals significant differences between the prime gaps and the spacings of the Zeta function's non-trivial zeros. With a KS statistic of 0.687 and a p-value of approximately 1.41×10^{-22}, the large KS statistic indicates a notable difference between the two distributions. The very small p-value strongly suggests rejecting the null hypothesis that the distributions are identical, confirming their statistical dissimilarity.

The results indicate no straightforward correlation between prime gaps and the Zeta function's non-trivial zeros. However, the numerical and sequential nature of this data suggests potential for uncovering more complex relationships. The aim is to identify patterns or correlations that are not immediately obvious, necessitating models that detect subtle, nonlinear connections. This exploration will be the next focus.

Various models were investigated, including unsupervised learning methods like principal component analysis and *t*-distributed stochastic neighbor embedding, along with cluster analysis. Supervised learning techniques such as polynomial regression, random forests, and gradient boosting machines were also explored. Additionally, time series analysis methods were considered, including auto regressive integrated moving average (ARIMA) models andlong short-term memory (LSTM) networks, as well as deep learning approaches such as neural networks.Lastly, the focus was on causal inference followed by an exploration of causal networks.

Principal component analysis (PCA) and *t*-distributed stochastic neighbor embedding (*t*-SNE) are effective for reducing dimensions and visualizing high-dimensional data structures. They aid in identifying clusters or patterns. Cluster analysis methods, such as *K*-means and hierarchical clustering, group data points to highlight patterns or anomalies. The *K*-means method identifies clusters in data by partitioning it into *K* distinct groups based on similarity, which is crucial for analyzing complex datasets but requires careful selection of *K* and can be computationally intensive. Applying these techniques to zeros or prime gaps could uncover significant groupings or distributions.

When a functional relationship is suspected, regression models, such as polynomial regression, are valuable for modeling and predicting variables. Polynomial regression is especially effective in capturing nonlinear relationships.

Random forests and gradient boosting machines, as ensemble methods, excel in identifying complex relationships and interactions among variables. They are also beneficial for analyzing feature importance, offering insights into which features, like certain prime gaps or zeros, are the most predictive.

Considering the data's sequential nature, time series analysis methods are applicable. Models such as ARIMA or LSTM neural networks can be used to analyze temporal dependencies in the data. The ARIMA model, standing for AutoRegressive Integrated Moving Average, is a time series forecasting method that uses differences of observations (integration) and lagged observations (autoregression) along with lagged forecast errors (moving average) to predict future values.LSTM neural networks, or Long Short-Term Memory networks, are a type of recurrent neural network (RNN) designed to remember long-term dependencies by incorporating memory cells that regulate the flow of information.

A multi-layered neural network is capable of identifying complex, nonlinear relationships within data. With an appropriate structure and enough training data, it has the potential to reveal underlying patterns in the distribution of prime gaps and zeros.

Causal inference aims to identify if there is a cause-effect link between variables. This differs from correlation or predictive modeling, which only show associations or forecast outcomes. Causal inference delves into understanding how one variable directly impacts another.Causal networks, unlike Bayesian networks, specifically model and uncover causal relationships in data. While Bayesian networks are for probabilistic inference and represent various dependencies, causal networks concentrate on cause-and-effect connections. They use causal inference theory to identify not just the relationships between variables, but also how these relationships influence each other causally.

3. Results

PCA and *t*-SNE are techniques for reducing dimensions and visualizing data, revealing structures in high-dimensional data. However, the data (prime gaps and non-trivial zeros of the Riemann Zeta function) are single-dimensional, making these techniques less applicable. PCA requires fewer components than the minimum of samples or features, and since the data has only one feature, reducing it to two components is not feasible. *t*-SNE could be applied to a smaller data subset due to its complexity and computational demands, but this approach is not ideal. Therefore, these methods were considered unsuitable for the data.

Cluster analysisgroups data based on similarities. This was applied to prime gaps and Riemann Zeta function's non-trivial zeros datasets using K-means clustering.This method's key parameter is the number of clusters (K). Choosing K is critical, and the Elbow method is used to determine an optimal K. Due to K-means' high computational demands, analysis is limited to a smaller range of K values and a sample size of 50 from each dataset is used. The findings, displayed in Figures 5 and 6, show these samples clustered into three groups. The plots visually represent each data point, color-coded by its assigned cluster.

The prime gaps sample clusters (Figure 5) exhibit differentiation, but their meaning and importance need more analysis. Likewise, the non-trivial zeros of the Zeta function (Figure 6) form clear clusters. Yet, understanding these clusters' relevance to the Riemann Hypothesis and their mathematical significance requires thorough examination. The clustering outcomes depend on the chosen sample and K-means algorithm's randomness, particularly during initialization. Hence, these clusters should be viewed as a part of an exploratory analysis rather than definitive findings. To gain deeper understanding, further studies using larger samples or different clustering methods are advisable. Also, examining the mathematical or theoretical implications of these clusters is crucial to connect the findings with prime number theory and the Riemann Hypothesis.

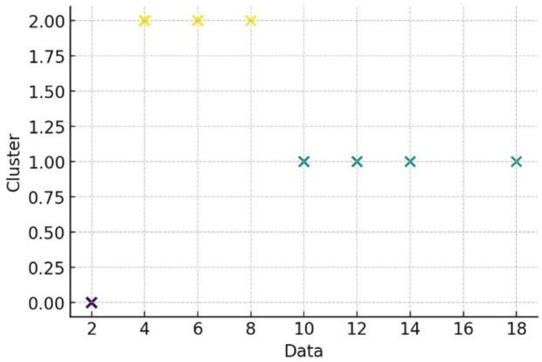

Figure 5. Clusters of sampled prime gaps.

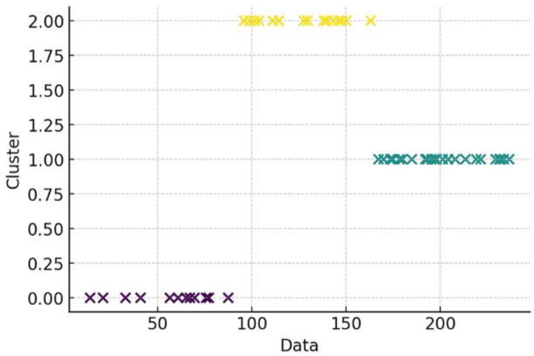

Figure 6. Clusters of sampled non-trivial zeros of the Zeta function.

A degree-3 polynomial regression model to explore nonlinear relationships is used. Prime gaps were considered as the target variable and non-trivial zeros of the Zeta function as the predictor. The model showed a mean squared error (MSE) of 8.19 and a negative R^2 value of −1.26. The MSE is a measure of the average squared difference between the observed actual outcomes and the outcomes predicted by the model. The optimal value of MSE is 0, indicating perfect predictions with no errors, while higher values indicate worse predictive accuracy, signifying a model's poor fit to the data. The high MSE indicates significant prediction errors. The R^2 value indicates the proportion of the variance in the dependent variable that is predictable from the independent variable. A negative R^2 value, normally close to 1 in good models, suggests the model fits worse than a basic mean line. This implies the model poorly represents the relationship between the Riemann Zeta function's non-trivial zeros and prime gaps. The negative R^2 and high MSE hint that the relationship, if present, is complex and not well-captured by a simple degree-3 polynomial model.

Random forest regression was also employed to analyze the relationship between the non-trivial zeros of the Riemann Zeta function and prime gaps, using the same data. The trained random forest model yielded an MSE of 17.57 and a negative R^2 of –3.84. This high MSE indicates substantial deviation in the model's predictions from actual values. The negative R^2 value reflects a very poor fit, suggesting the model does not effectively capture the data's relationship.

Before using a gradient boosting machine (GBM) model on the data (prime gaps and non-trivial zeros of the Riemann Zeta function), the analysis goal needs to be defined. GBM models suit supervised learning, like regression (predicting a continuous value) or classification (predicting a discrete category). In this case, GBM for regression can be applied, attempting to predict prime gap size from a non-trivial zero of the Zeta function, or the reverse. This assumes a potential functional relationship between the two variables.

Due to the unequal number of data points between zeros and prime gaps, the data were paired before analysis. The data were divided into a training set for model development and a testing set for validation. A GBM regressor was employed to model the relationship between the zeros and prime gaps. The model's performance was then assessed on the test set using the metrics of MSE and R^2. After training and evaluating the GBM model, an MSE of 16.37 and a negative R^2 value of –3.51 were observed. The high MSE indicates the model's predictions are far from the actual values. The negative R^2 suggests the model performs worse than a basic horizontal line (the target variable's mean). Thus, the GBM model analysis fails to uncover a predictive relationship between the Riemann Zeta function's non-trivial zeros and prime gaps.

To analyze the data using time series methods, the prime gaps and non-trivial zeros of the Riemann Zeta function were treated as sequences in time order. The ARIMA model for forecasting was used, recognizing its effectiveness in capturing temporal patterns in data. While LSTM neural networks are potent for such analysis, they necessitate a deep learning framework like TensorFlow, not utilized due to its complexity.

ARIMA with parameters (p, d, q) were used to analyze both prime gaps and non-trivial zeros of the Riemann Zeta function, initially setting them at common values (ARIMA(1, 1, 1)) and adjusting based on initial results.The ARIMA model's parameters (p, d, q) represent the order of the autoregressive part (p), the degree of differencing needed to make the time series stationary (d), and the order of the moving average part (q), respectively, which are crucial for specifying the model's structure and capturing the time series' characteristics. For prime gaps, ARIMA parameters (AR = –0.1974, MA = –0.9263) revealed autocorrelation. Tests (Ljung-Box, Jarque-Bera) confirmed that residuals were like white noise and normally distributed. Similarly, for the Zeta function's non-trivial zeros, parameters (AR = 0.9990, MA= –0.8527) suggested autocorrelation. Residuals, per Ljung-Box and Jarque-Bera tests, were also akin to white noise and normally distributed.Although the ARIMA models reveal certain statistical aspects of the data, they do not clarify the connection between prime gaps and the non-trivial zeros of the Zeta function, particularly in relation to the Riemann Hypothesis. The analysis is more a statistical examination of each dataset than a probe into their deeper mathematical or temporal interrelations.

A basic neural network model was attempted to use, similar to the previous regression approach, to predict prime gap sizes from non-trivial zeros of the Zeta function, using the same data as in the GBM model. This neural network would have a few layers and be trained on the training dataset. However, this approach was abandoned because it required TensorFlow, a key library for neural networks, which was not used. To proceed with TensorFlow was discarded, anticipating that neural networks would encounter the same issues as the GBM model.

In this study of the relationship between prime gaps and non-trivial zeros of the Riemann Zeta function, causal inference would mean determining if changes in one variable directly cause changes in the other. This is challenging for several reasons. First, causal inference needs a strong theoretical or empirical basis for a causal link, which does not exist between these variables. Second, it typically relies on experimental or longitudinal data showing interventions or changes over time, which the observational dataset lacks. Lastly, methods like Granger causality tests, instrumental variable analysis, or counterfactual models [16,17] require specific conditions not met by the data. Despite these challenges, a basic causal analysis using Granger causality can be conducted, but the results should be interpreted with caution and may not conclusively prove causality.

The Granger causality test is used to check if one time series can predict another. Though it focuses more on prediction than actual causation, it is a common first step in causal analysis. This test was applied to see if the non-trivial zeros of the Riemann Zeta function can predict the prime gaps, and vice versa. The test showed statistical significance (p-values less than 0.05) at all tested lags (1, 2, 3), indicating that past values of non-trivial zeros might predict future prime gaps. However, these results, while statistically significant, do not confirm a true cause-and-effect relationship but suggest a potential predictive link between the datasets. This is a promising step in this exploration, leading to the examinationof causal networks next.

To build a causal network, pinpointing possible causal linksis a convenient start, drawing on the prior knowledge of Granger causality test outcomes. For examining causal networks related to prime gaps and non-trivial zeros of the Riemann Zeta function, potential causal connections must be hypothesized. Considering the data and mathematical context, these assumptions are entirely speculative. Based on the Granger causality tests that suggested a predictive link between prime gaps and non-trivial zeros of the

Riemann Zeta function, two assumptions for the causal network analysis are proposed: 1) Characteristics of prime gaps may causally affect subsequent non-trivial zeros. 2) Properties of non-trivial zeros could influence the characteristics of subsequent prime gaps.

In this analysis of prime gaps and non-trivial zeros of the Riemann Zeta function, regression-based methods were used to estimate causal effects, treating one dataset as the predictor and the other as the outcome. For model A (prime gaps predicting non-trivial zeros), an MSE of 7.65 and an R^2 of −1.11 were found, indicating a poor fit. The negative R^2suggests the model fails to accurately represent the relationship. Model B (non-trivial zeros predicting prime gaps) showed an MSE of 4080.65 and a low R^2 of 0.049, implying slight predictive ability but overall weak model fit and inaccurate predictions. These results indicate that the hypothesized causal relationships, modeled through linear regression, are not strongly supported by the data.

Summarizing this investigation into the relationship between prime gaps and non-trivial zeros of the Riemann Zeta function, three key findings emerge: 1) The Granger causality tests indicate a predictive relationship between prime gaps and non-trivial zeros. 2) The relationship is not simply linear, as shown by regression and polynomial model results. 3) A complex, possibly nonlinear relationship, maybe with feedback loops, is suggested by the causal network analysis. Moving forward, further investigations can be based on these three outcomes.

Considering the above results as assumptions, a model to study the link between prime gaps (*X*) and the non-trivial zeros of the Riemann Zeta function (*Y*) is suggested. The model's equation is $Y = aX = bX^2 + cX^3 + \varepsilon$, where *a*, *b*, and *c* are coefficients showing how *X* affects *Y*, and *ε* represents unexplained variability. This aims to explore the complex, nonlinear connection between *X* and *Y*. The model's performance is indicated by an MSE of 4033.34 and a mean absolute error (MAE) of 55.58. The MSE and MAE values help assess the polynomial model's fit for the intricate link between prime gaps and the Riemann Zeta function's non-trivial zeros. However, determining if these MSE and MAE values are high or low is challenging without context from number theory. Nevertheless, Figure 7 provides a visual comparison of actual data against the model's predictions.

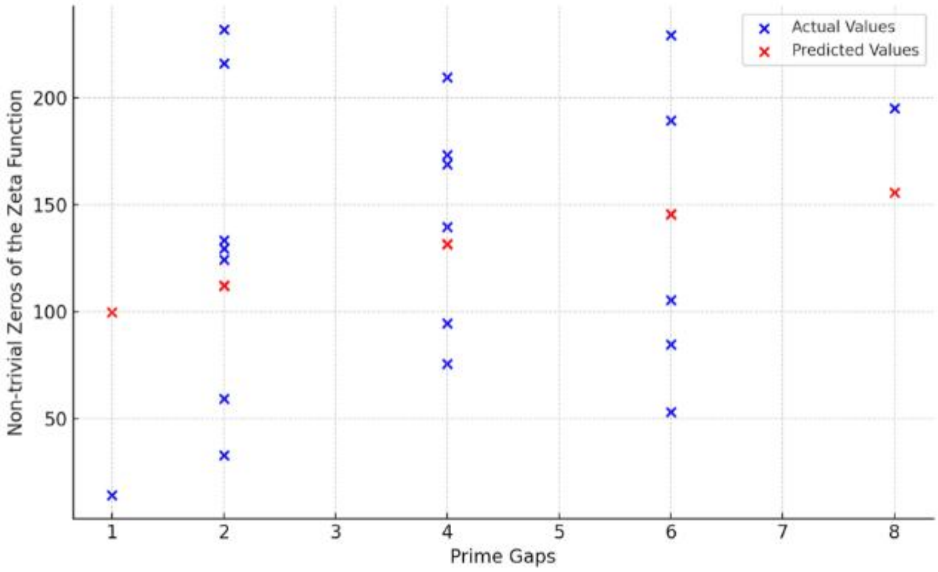

Figure 7. Polynomial regression: prime gaps vs. non-trivial zeros.

4. Discussion

This paper adopts a new approach to the Riemann Hypothesis, examining prime gaps and non-trivial zeros of the Riemann Zeta function using statistical models and experimental analysis. Challenges include unclear causal connections and complex relationship modeling. Improvements suggested are: 1) Linking experimental findings more closely with theoretical aspects of the Riemann Hypothesis. 2) Enhancing statistical models for better analysis of complex relationships. 3) Deepening investigation into causal links. 4) Integrating findings more thoroughly with existing Riemann Hypothesis research and theories. To address the latter concern, some recent work merits consideration.

Tegetmeyer'spaper [14] attempts to prove the Riemann Hypothesis using the Zeta function's integral representations and complex analysis. One can enhance this paper's study by integrating his approaches: 1) Applying integral representations of the Zeta function, 2) Using his methodology to solve these integrals for the Zeta function's real and imaginary parts, and 3) Identifying conditions for the Zeta function's zeroes. This integration could deepen the mathematical understanding in this study's exploration, especially concerning the Zeta function's behavior in the Riemann Hypothesis context, linking this paper's experimental observations with formal proofs.

Remmen's paper [15] links scattering amplitudes in physics to the Riemann Zeta function, constructing an amplitude reflective of the Zeta function's properties. It maps scattering amplitudes' physical attributes to the Zeta function and establishes a closed-form amplitude. The paper also relates locality and unitarity with the Riemann Hypothesis. Integrating these concepts into this paper's work could offer a novel view, connecting physical phenomena with the Zeta function's mathematical traits.

Borwein et al.'s book [18] offers an extensive review of the Riemann Hypothesis, covering its history, core mathematics, and diverse proof attempts. Its integration into this paper's study could enhance understanding by providing: 1) In-depth historical insights into the hypothesis' importance, 2) Advanced mathematical proofs and concepts, and 3) Analysis of various proof strategies, both successful and not. This inclusion would enhance the historical and theoretical components of this paper's investigation, providing a more complete picture of the Riemann Hypothesis and possibly offering new avenues for investigation.

Levinson's paper [19] offers an in-depth analysis of the Riemann Zeta function's zeros, especially on the critical line. Its mathematical techniques and findings could enhance this study of the Riemann Hypothesis. Key aspects include: 1) Techniques for estimating zero counts on the critical line using complex analysis, 2) Discussions on zero density and its impact on the Hypothesis, and 3) Detailed proofs and calculations that could provide a solid mathematical base or comparison for the present exploration. Integrating these elements could deepen the mathematical understanding, reinforce the analysis, and potentially reveal new insights related to the Hypothesis.

Conrey and Ghosh's work [20] analyzes the Riemann Zeta function's mean values and zeros, particularly along the critical line. Their paper's contributions to this research could include: 1) Techniques for estimating the Zeta function's mean values, 2) In-depth analysis of its zeros, especially on the critical line, and 3) The employed mathematical methods. These elements could significantly enrich this study of the Riemann Hypothesis by deepening the understanding of the Zeta function's behavior.

Zhang's [21] study offers a novel absolute convergent formulation of $\zeta(s)$ derived from the Hadamard product. The work also addresses the features of Riemann Zeta function non-trivial zeros, their multiplicity, and their arrangement in increasing order. It also examines the functional equation $\zeta(s) = \zeta(1 - s)$ in depth, leading to the conclusion that the Riemann Hypothesis is correct based on the features of the zeros. These ideas could be incorporated into this paper's investigation, particularly in respect to the distribution and properties of the Riemann Zeta function's non-trivial zeros.

In Abe's study [22], a compelling point about the Riemann Zeta function is made: If only a finite number of nontrivial zeros fall outside the critical line, then all nontrivial zeros must indeed be on that line. This could offer valuable insights for this paper's research on the Riemann Zeta function's non-trivial zeros.

Nayebi's paper [23] explores how the Riemann Hypothesis could be represented as a Diophantine equation, allowing for computational analysis. This approach links the hypothesis to Hilbert's Tenth Problem, furthering its understanding through number theory and computational mathematics. This perspective could add a novel computational dimension to this paper's study, particularly in understanding the Riemann Zeta function's zeros through numerical and algorithmic methods.

The present study provides intriguing and novel insights, but needs considerable refinement to solidify its contributions to the Riemann Hypothesis along these lines.However, the purpose of offering insights into a mathematical proof or disproof of the Hypothesis was met.

To enhance this study on the Riemann Hypothesis, future research should deepen the link between this paper's experimental results and the Riemann Hypothesis. Refining statistical models is crucial to better understand the complex relationships between prime gaps and non-trivial zeros. Investigating causal connections between such variables is also vital. Additionally, integrating this paper's findings with existing literature will enrich the study's context, thereby contributing more effectively to the Riemann Hypothesis discourse. These improvements will significantly strengthen this paper's suggested methodology on understanding this difficult mathematical problem.

5. Conclusion

The exploratory analysis of prime gaps and non-trivial zeros of the Riemann Zeta function, using various statistical models, suggests several potential approaches to the Riemann Hypothesis. The polynomial regression model indicates a complex, possibly nonlinear relationship between these entities, demonstrating the inadequacy of simple linear models for such intricate mathematical concepts.

The Granger causality tests hint at a predictive relationship between prime gaps and non-trivial zeros, although this does not necessarily imply direct causation. This finding opens up possibilities for further investigation into underlying data patterns.

However, the attempts at applying causal inference models faced significant challenges, highlighting the difficulties of establishing causal links in number theory. This study represents a first exploratory step towards comprehending the connection between prime gaps and non-trivial zeros, and it implies that future research should incorporate more sophisticated mathematical techniques with statistical and experimental methodologies.

In summary, this analysis offers intriguing insights and directions for further study but also emphasizes the Riemann Hypothesis's complexity, calling for a comprehensive approach that melds rigorous mathematics with sophisticated statistics and experimental methods. Progressing towards solving or advancing significantly on the Riemann Hypothesis remains a formidable task in mathematics.

References

1. Matsushita, R.; Da Silva, S. A power law governing prime gaps. *Open Access Libr.*,**2016**, *3*, e2989.
2. Hardy, G.H.; Littlewood, J.E. Some problems of "partitionumerorum". III. On the expression of a number as a sum of primes. *Acta Math–Jursholm*, **1923**, *44*, 1–70.
3. Lemke Oliver, R.J.; Soundararajan. K. Unexpected biases in the distribution of consecutive primes. *Proc. Natl. Acad. Sci. U.S.A.***2016**, *113*, E4446–E4454.
4. Dahmen, S.R.; Prado, S.D.; Stuermer-Daitx, T. Similarity in the statistics of prime numbers. *Physica A*, 2001, *296*, 523–528.
5. Ares, S.; Castro, M. Hidden structure in the randomness of the prime number sequence? *Physica A*, **2006**, *360*, 285–296.
6. Szpiro, G.G. Peaks and gaps: spectral analysis of the intervals between prime numbers. *Physica A*, **2007**, *384*, 291–296.
7. Wolf, M. Applications of statistical mechanics in number theory. *Physica A*, **1999**, *274*, 149–157.
8. Kumar, P.; Ivanov, P.C.; Stanley, H.E. Information entropy and correlations in prime numbers. arXiv: cond-mat/0303110, 2003.
9. Wolf, M. 1/f noise in the distribution of prime numbers. *Physica A*, **1997**, *241*, 493–499.
10. Chandra, A.K.; Dasgupta, S. A small world network of prime numbers. *Physica A*, **2005**, *357*, 436–446.
11. Gamba, Z.; Hernando, J.; Romanelli, L. Are prime numbers regularly ordered? *Phys. Lett. A*, **1990**, *145*, 106–108.
12. Lattanzi, D. Statistical distributions of prime number gaps. *J. Adv. Math. Com. Sci.*, **2024**, *39*, 36–61.
13. Mayer, B.; Monteiro, L.H.A. A numerical study on the regularity of *d*-primes via informational entropy and visibility algorithms. *Complexity*, **2020**, 1480890.
14.Tegetmeyer, B. Proof of the Riemann Hypothesis. *ArXiv*, **2023**, https://arxiv.org/pdf/2201.06601.pdf.
15. Remmen, G.N. Amplitudes and the Riemann Zeta function. *Phys. Rev. Lett.*, **2021**, *127*, 241602.
16. Cunningham, S. Causal inference: The mixtape. New Haven: Yale University Press, 2021.
17.Chen, B.; Pearl, J. Regression and causation: A critical examination of econometrics textbooks. *Real-World Econ. Rev.*, **2013**, *65*, 2–20.
18. Borwein, P.; Choi, S.; Rooney, B.; Weirathmueller, A. The Riemann hypothesis: A resource for the afficionado and virtuoso alike. Berlin: Springer, 2007.

19. Levinson, N. More than one third of the zeros of Riemann's Zeta-function are on σ = 1/2. *Adv. Math.*, **1974**,*13*, 383–436.
20. Conrey, J.B.; Ghosh, A. On mean values of the Zeta-function. *Mathematika*, **1984**, *31*, 159–161.
21. Zhang, W.A Proof of the Riemann hypothesis based on anew expression of the completed Zeta function. Preprints.org, 2023.
https://doi.org/10.20944/preprints202108.0146.v25
22. Abe, Y. A remark on the Riemann zeta function. Cambridge Open Engage, 2023. https://doi.org/10.33774/coe-2022-d3nj5-v9
23.Nayebi, A. On the Riemann hypothesis and Hilbert's tenth problem. arXiv:2311.09464, 2023.