Testing and understanding the jackknife approach to mutual information estimation

Grishma Mainali

22101489

A report submitted for

INFO7016 Postgraduate Project A

in partial fulfillment of the requirements for the degree of

Master of Data Science

Supervisor: **Prof. Paul Hurley**

School of Computer, Data and Mathematical Sciences

Western Sydney University

October 2024

ABSTRACT

Mutual information (MI) is a fundamental notion in information and probability theory, and it is extensively employed as a measure of the relationship between two random variables. However, present MI estimate methods tend to be affected by bias and variance in the dataset. This research project focusses on the jackknife resampling method as an effective approach for improving MI estimates by minimizing bias and variance. This study is currently in the planning stage and will involve studying relevant literature, developing algorithms, and preparing for implementation using Python programming. For the implementation part, the jackknife approach will be evaluated under various scenarios using synthetic datasets with controlled parameters such as sample size, noise levels, and dependence patterns. The goal is to assess the jackknife method's performance on synthetic datasets and compare it with other commonly used MI estimation techniques. The expected outcome of this study will evaluate if the jackknife approach produces more precise and reliable estimations of MI, especially with small data sample. The outcomes of this study will aim to contribute to the development of more effective MI estimate techniques, providing useful insights for both theoretical research and practical applications.

ACKNOWLEDGEMENT

I'd want to use this opportunity to offer my heartfelt gratitude to my supervisor, Professor Paul Hurley, for his constant guidance and valuable support throughout this research. His insights and encouragement helped shape my work.

I am deeply grateful to my parents for everything they have done for me. Their unending love, sacrifices, and unflinching support served as the foundation for my journey. Thank you, dear sister, for always being yourself, for your love, and for being a constant source of joy and inspiration. I also want to express my heartfelt gratitude to my other family members, whose love and encouragement have always encouraged me.

Finally, I'd want to thank my friends for their encouragement and companionship, which have made this journey even more significant.

Table of Contents

Chapter I – Introduction

Mutual information (MI) can be defined as a general measure of relationship between two random variables. Simply put, it is the amount of information obtained about one random variable by observing another random variable (Wikipedia, 2024). It is an important tool in a variety of domains, including data analysis, machine learning, communication systems, and neurology. MI provides insights into the underlying relationships between data points by calculating how much information one variable provides about another, making it fundamental in figuring out dependencies within complex systems. MI is closely related to the concept of entropy as it can also be defined as reducing the uncertainty of one random variable when another is known. As a result, a high mutual information value suggests a significant reduction in uncertainty, while a low value indicates a minor reduction. If the mutual information is 0, it indicates that the two random variables are independent (Aznar, 2021).

Even while MI has been receiving more attention and study recently, its history and relevance are strongly tied to the groundbreaking work of the renowned Claude Shannon, who introduced the measurement of Entropy for quantifying information (Zeng, 2015). MI extends Shannon's entropy, providing a broader measure that can capture shared information between two variables. MI has been studied and reinvented over time, to make it one of the critical measures for information shared between two random variables. However, measuring the MI accurately and effectively from data is difficult since estimate methodologies include bias and excessive variance (Shannon, 1948). Many of these methods have a major shortcoming: they rely on tuning parameters, which add instability and might result in incorrect estimates when dealing with various types of data distributions or sample sizes.

The motivation for choosing "Testing and Understanding the Jackknife Approach to Mutual Information Estimation" as the research topic originates from an urge to improve existing MI estimation methodologies. The jackknife approach (Zeng, et al., 2018) offers a viable solution by utilizing Jackknife resampling approach to reduce bias and deliver reliable variance estimates. According to the above stated research paper, the jackknife approach offers considerable enhancements and is effective for MI estimation. But this method has not yet been properly tested, applied, or sufficiently explored to show its efficacy. Furthermore, there is not enough comprehensive research that compare the performance of the jackknife method to other widely used MI estimate techniques. The lack of evidence provides a unique opportunity to contribute to the area by not only implementing the jackknife approach but also extensively testing its performance. Thus, the purpose of this research project is to understand, implement, and evaluate the jackknife technique to MI estimation using Python, as well as to contribute further insights into this field of study.

Problem Statement

Accurate MI estimation is important for understanding the relationships between variables in many domains. However, present estimating approaches are frequently constrained by bias and excessive variance, resulting in unsatisfactory performance. Other MI estimation methods are also vulnerable to the tuning parameters used, which might result to estimate instability. The issue needs the investigation of alternate estimating methods that provide more accuracy and stability without the need for parameter adjustment.

The jackknife approach has been presented as an alternate solution that may solve these limitations. The jackknife method, which employs resampling techniques, aims to eliminate bias and produce more dependable variance estimates, resulting in more accurate MI estimations. However, the method's effectiveness has yet to be thoroughly validated across various contexts, and there is a not enough of extensive study comparing its performance to other regularly used MI estimate strategies.

Research Objective

- Understanding the Jackknife Method: Learn about the jackknife strategy to mutual information estimation, as proposed in the referenced paper.
- Implementation: Implement the jackknife MI estimate technique in Python programming language.
- Evaluation: Evaluate the jackknife MI estimator's performance on data and evaluate the performance it can generate.
- Documentation: Provide documentation of the implementation of the complete method as well as a detailed report on the findings.

Hypothesis

Other commonly used methods for calculating MI have been shown to have unstable statistical performance due to the need for a set of tuning parameters, making them unable to deliver the optimal estimate. For estimating MI, the Jackknife method offers measurements of MI that are more precise, dependable, and stable as this method does not require to work with tuning any parameters (Zeng, et al., 2018).

Research Question

How does the jackknife approach's accuracy and bias compare with other mutual information estimating techniques?

This study could provide substantial results since it addresses a major gap in our current understanding of MI estimate. This work seeks to provide useful insights into the performance and utility of the jackknife methodology in comparison to established approaches by conducting a complete analysis that involves literature review, algorithm development, and testing on synthetic data. If proven effective, the jackknife method could be a useful tool for estimating MI, adding to fields such as machine learning, data science, and communication theory.

Chapter II – Literature Review

A good literature assessment is critical for setting the correct foundation for a research study. This portion of the report includes a comprehensive review of the underlying concepts and available research on Mutual Information (MI) estimation and the Jackknife resampling technique. It assesses relevant research works on MI estimation methods, resampling approaches, as well as associated theories in the broader field of information theory which would help in identifying gaps in the literature that this research aims to address.

Mutual Information

In probability and information theory, Mutual Information (MI) of two random variables is a measure of their dependence upon one another. Mutual information is fundamentally linked to the concept of entropy of a random variable, which is a fundamental term in information theory that measures the predicted "amount of information" stored in a random variable (Wikipedia, 2024). Correlation coefficient, a well-known concept of identifying relationships between variables, is limited only for linear dependencies. But the concept of MI can be applicable even for non-linear dependencies which makes it so much more crucial. Claude Shannon defined and analyzed the amount of shared information in his influential work "A Mathematical Theory of Communication," however he did not use the term "mutual information." Robert Fano coined this word later (Wikipedia, 2024).

The foundation for understanding the fundamental ideas of mutual information and entropy in the literature review carried out for this research study remains Claude Shannon's pioneering work. Shannon established the notion of entropy, which is a mathematical tool for quantifying the uncertainty in information. Shannon Entropy, named for Claude Shannon, is a type of information entropy that measures the uncertainty or randomness of information content. The higher the entropy of information, the more information it contains and the more effective it is at reducing uncertainty (Zhang, 2023).

Built upon the Shannon's foundations, "Elements of Information Theory" is another important work of literature that is strongly followed and referenced for gaining knowledge and conducting in-depth study on mutual information and different MI estimation strategies. This study is an important resource for identifying the complexities of MI estimate and its implications in various areas, since it offers a comprehensive grasp of both the theoretical and practical elements of MI (Cover & Thomas, 2006).

Mathematically, Mutual Information between two random variables, say X and Y, is calculated as:

$$I(X;Y) = H(X) + H(Y) - H(X,Y)$$

Where the entropies of random variables X and Y are denoted by H(X) and H(Y), respectively, while H(X,Y) represents their joint entropy. MI can range from zero, implying absolute independence between X and Y, to higher levels, suggesting greater dependency. The Venn diagram below is a great reference for graphically understanding MI's features. The two circles reflect the uncertainty of variables X and Y, while the union represents the combined Shannon entropy. The intersection also represents the mutual information. Observing Y, or reducing the blue circle indicating Y's uncertainty, reduces the uncertainty about X (the red circle) by an amount equal to the mutual information (the intersection). The mutual information is symmetric, so observing X instead of Y would reduce uncertainty about Y by the same amount. Furthermore, we can extract that when the two variables X and Y are totally independent (i.e., there is no overlapping between the two circles), the mutual information is zero (Crosato & Morris, 2023).

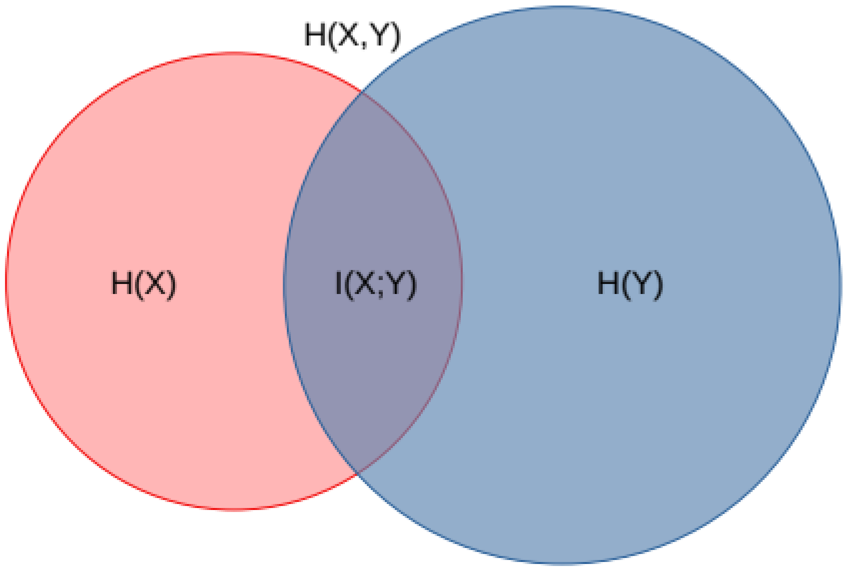

Figure 1: Venn Diagram representation of MI (Crosato & Morris, 2023)

Challenges in MI Estimation

Several reasons contribute to the difficulty of effectively estimating MI, including the complexity of entropy calculation and the sensitivity of MI estimations to data noise and sample size.

There have been made many efforts to introduce methods and techniques which can produce MI estimates with higher accuracy. Early approaches, such as Fraser and Swinney's (1986) histogram-based estimator, divided data into discrete bins to estimate probabilities. This strategy, however, fosters significant bias and variance, particularly in small datasets, due to random binning of continuous data. (Moon, et al., 1995) suggested a more advanced method utilizing kernel density estimation (KDE), which smoothes the data to improve MI estimation. Though KDE outperforms histograms, it is computationally expensive, particularly for high-dimensional data, and requires careful selection of bandwidth parameters. K-nearest neighbor (KNN) estimator which was introduced in 2004 is a non-parametric approach that seeks to optimize the bias-variance tradeoff by computing MI using local densities surrounding data points. The KNN estimator's appeal stems from its adaptability in high-dimensional domains. However, when the dataset size expands, the computational cost increases, and the accuracy is affected by the k parameter. (Kraskov, et al., 2004). Recently, neural network-based approaches have been created, however they have not been able to eliminate the prior issues, as they frequently suffer from training instability and bias (Khan, et al., 2007).

Across all these methods, issues such as hyperparameter sensitivity, high computational costs, and poor performance in small datasets persist, prompting the inclusion of resampling techniques such as the jackknife approach, which aims to reduce bias and variance without requiring complex tuning parameters.

Resampling Techniques

Resampling is a technique for repeatedly taking samples from the training dataset. These samples are then utilized for remodeling a specific model to obtain further information about the fitted model. The goal is to learn more about a sample, enhance its accuracy, and estimate its uncertainty. Resampling approaches, such as jackknife and bootstrap, are commonly employed in statistical analysis to eliminate bias and estimate variance. These methods correspond to repeatedly gathering samples from the dataset and recalculating the statistic of interest, which allows to examine its variability and obtain more reliable estimations (Arya, 2023).

As mentioned above, Bootstrap and Jackknife are the two most utilized resampling methods in the field of data and statistics. Efron (1979) invented the bootstrap approach, which includes producing repeated resamples of the original dataset using sampling with replacement and computing the statistic of interest for each resample (Wikipedia, 2024). This method enables the estimate of confidence intervals and other statistical features, making it a prominent technique in domains like as machine learning, data science and econometrics. Bootstrapping does not generate new data; rather, it uses the original sample as a proxy for the true population and draws random samples from it. As a result, the primary assumption for bootstrapping is that the original sample accurately represents the actual population. The various combinations of values in the simulated samples provide a general indication of the variability between random samples selected from the same population (Frost, 2024).

The jackknife approach, first introduced by Maurice Quenouille in 1949 and later improved by John Tukey, works by systematically removing one observation from the dataset at a time and recalculating the intended statistic (Wikipedia, 2024). The jackknife method is especially good in reducing bias and offers a simple way to estimate standard errors. The Jackknife's primary application is to eliminate bias and analyze variance for estimators. Using the jackknife method specifically to MI estimation aims to eliminate the bias inherent in standard methods, resulting in more trustworthy results, especially in conditions with small sample sizes or excessive noise (Glen, 2019).

Jackknife Resampling for MI

In their study "Jackknife Approach to the Estimation of Mutual Information," published in 2018, Xianli Zeng, Yingcun Xia, and Howell Tong proposed an inventive technique for calculating MI. This is the main work which will be referenced for this research study as we implement and access the jackknife approach. The resampling approach called the "jackknife approach" attempts to lessen the bias and variation seen in conventional MI estimation techniques. The authors give a thorough mathematical explanation of the jackknife method and show how well it works with complex data distributions and small sample numbers. The study presents preliminary findings that demonstrate the efficiency of the procedure in comparison to traditional techniques (Zeng, et al., 2018). While Zeng et al.'s research established the method, it has not been extensively adopted or developed.

This study expands on Zeng et al.'s fundamental work by implementing and testing the jackknife approach for MI estimation in a variety of circumstances, including differing sample sizes, degrees of dependency between variables, and noise levels. Furthermore, although earlier research has primarily concentrated on the theoretical and mathematical components of MI estimation, this study emphasizes the practical use of the jackknife approach in Python.

Data Sources

In prior works on MI estimation, researchers primarily used synthetic datasets, which allow for controlled trials and facilitate the testing of MI estimation algorithms under the best possible circumstances. Synthetic data allows for the manipulation of variables such as sample size, noise, and degree of dependency, providing a clear picture of the strengths and shortcomings of various estimate methods.

Similarly, this study will use synthetic datasets to assess the efficiency of the jackknife approach under a variety of situations. This strategy provides exact control over critical variables, ensuring that the results truly reflect the method's performance.

Chapter III: Methodology

The fundamental goal of this research is to comprehend, apply, and assess the jackknife technique to Mutual Information (MI) estimate. This chapter describes the methodology used to attain these aims, including the techniques, tools, and strategies used throughout the research. The general purpose is to provide a strong foundation for testing the jackknife approach to estimating MI, comparing it to other methods, and assessing how it performs using a variety of performance criteria such as accuracy, bias, and computational efficiency. The process also includes creating experiments with synthetic datasets to control certain variables such as sample size and noise levels.

As stated in the introduction section of this report, the specific goals of this research project are to understand the use of the jackknife resampling approach in the calculation of Mutual Information, to then put the theoretical aspects into practice using Python programming, to evaluate the results, and to properly document the entire process to ensure reproducibility and clarity for future references.

Research Approach

The study approach used here is separated into four phases: literature review, implementation, experimentation, and assessment. These phases are planned to fulfil the project's unique goals and to give a clear path towards addressing the research questions.

i. Literature Review: As described in Chapter II, a detailed literature analysis is being carried out to identify obstacles in mutual information estimates and evaluate existing approaches. The study process will continue as the project progresses through its various implementation phases. The review presents the theoretical foundation required to implement the jackknife approach and highlights the shortcomings in current approaches, which motivated the use of the jackknife technique.
ii. Implementation: The jackknife approach for MI estimation will be implemented from scratch in Python. Synthetic datasets with regulated dependencies between random variables will be created to assess the accuracy of MI estimate. The jackknife resampling technique will be used to compute the MI between two variables. Along with the jackknife method, other MI estimate algorithms will be implemented or adapted from existing libraries for comparison.
iii. Experimentation: Several computational experiments will be carried out to assess the performance of the jackknife MI estimator. These could involve producing datasets with varied degrees of dependency between variables, introducing noise into the datasets to imitate real-world settings, and validating the estimator's accuracy and bias with small and big datasets.
iv. Assessment: Final step in the implementation phase would be to assess the performance of the implemented processes and the results that was achieved. The research will aim to look into the accuracy of the estimate obtained as well as the bias and variance in the estimates.

Python Programming

This research will use Python programming and will rely on various Python libraries, which are necessary for applying the methodology. NumPy is a Python library which would be used for numerical computations, array manipulation, and synthetic dataset generation. Similarly, SciPy would be utilized for scientific computing, specifically entropy and MI calculations. Pandas would be used for manipulating and processing structured data sets. Matplotlib/Seaborn could be used if required for visualizing results using graphs of MI estimates, bias, and variation. Scikit-learn contains certain in-built MI estimation which would be used for comparison purposes.

Chapter IV – Preliminary Results

As part of this research, some code was written to calculate mutual information (MI) between pairs of variables and to apply the jackknife resampling technique to eliminate bias in MI estimate. The major goal was to understand the differences in MI between dependent and independent variables, as well as to investigate how the jackknife method can increase the accuracy of MI estimates.

I. Mutual Information for Independent and Dependent Variables

The initial stage was to create two datasets, one with independent variables and the other with dependent variables that had some noise. The data was generated using NumPy library in Python. The mutual information for each dataset was calculated using a customized MI implementation based on probability distributions, and the results were compared to those produced by scikit-learn's mutual_info_score. The findings revealed that the MI for the independent dataset was near to zero (showing no dependency), whereas the dependent dataset had a much greater MI, indicating a link between the two variables.

II. Jackknife Resampling for Mutual Information

Next, the jackknife resampling technique was used on both the independent and dependent datasets to estimate MI with less bias. The jackknife method deleted one observation at a time, recalculating the MI and then averaging the results. This method gave a bias-corrected estimate of MI as well as a variance estimate for the jackknife samples. For both the independent and dependent datasets, the jackknife estimate was nearly identical to the direct MI estimates, with a little improvement in variance reduction.

For example, the direct MI estimate for the dependent variables was 1.5416, whereas the jackknife estimate was slightly higher (1.5415) with low variance. Similarly, for the independent variables, both the direct and jackknife estimates showed essentially minimal mutual information, as expected.

III. Comparison to Scikit-Learn's Mutual Information

The findings were also compared to scikit-learn's implementation of mutual information. Interestingly, scikit-learn regularly produced lower MI values than the direct and jackknife estimations. For example, the MI for the independent variables was 0.0051 using scikit-learn, but the direct estimate was 0.0074. The difference could be due to variations in the underlying techniques employed by scikit-learn and the custom MI implementation.

IV. Small Dataset Example and Bias Correction

An additional experiment was carried out on a small dataset with perfect dependency between variables to investigate the performance of the jackknife method in situations when MI is expected to be higher. The direct MI estimate for this dataset was 0.7219, and after using jackknife, the bias-corrected estimate climbed to 0.8111, with an estimated variance. This result demonstrates the jackknife method's capacity to account for bias in smaller datasets.

V. Linear Relationship and Correlation

In another experiment, a linear relationship between two continuous variables was simulated, with MI calculated using the direct, jackknife, and scikit-learn algorithms. The Pearson correlation coefficient and R^2 values were used to assess the data's linearity. The correlation coefficient of 0.9882 and R^2 value of 0.9765 indicate a strong linear relationship. Interestingly, the MI estimate from the jackknife approach was 8.0793, but the direct and scikit-learn methods returned lower values, highlighting the need of bias correction.

These initial findings show that the jackknife strategy is effective at estimating MI with low bias and variation, laying a strong platform for further investigation of MI estimation strategies. The trials also showed significant limitations in existing MI calculation algorithms, notably those included in popular libraries like scikit-learn, emphasizing the significance of bias correction in finite sample size circumstances. Moving on, the code will be modified to include more datasets and more complex dependencies to test the jackknife approach's robustness.

The code for these preliminary results has been uploaded to GitHub and can be accessed at the following link: https://github.com/IMGrishma17/Postgraduate_Project.git

Chapter V – Summary and Research Plan

This report lays the groundwork for future research and implementation by investigating the jackknife approach to Mutual Information (MI) estimation. The first goal was to examine the existing approaches for MI estimation, specifically the difficulties in obtaining accurate and unbiased estimates. A thorough literature analysis was carried out to discover existing techniques. The shortcomings in these approaches were studied, which included inclusion of bias and variance in estimation, computational inefficiency, and poor performance on small datasets. All these limitations observed through this research highlighted the importance of researching other methodologies such as the jackknife method.

Through a detailed study of (Zeng, et al., 2018) research on the jackknife methodology, the theoretical foundations of this method were thoroughly studied as well as its potential to solve fundamental drawbacks of existing MI estimate methods. Furthermore, a tentative implementation plan was developed, including the procedures needed for algorithm development, data generation, and performance evaluation. An initial draft of the methodology was also created, outlining a detailed plan for implementing the jackknife method, conducting experiments on synthetic datasets, and comparing performance to alternative estimate techniques. Overall, the purpose of this study is to lay the groundwork for further investigation and implementation in the subsequent phase of the research.

Research Plan for PPB

Postgraduate Project B will focus on the implementation, experimentation, and validation of the jackknife approach for MI estimation, building on the foundation given in this report, Postgraduate Project A. The major objectives for are listed below:

I. Implementation of MI and Jackknife Approach: The first stage will be to script the algorithm for MI calculation as well as MI estimation using Jackknife technique in Python. This will consist of developing functional algorithms, based on the theoretical concepts learned through above literature reviews. Python libraries such as NumPy, SciPy, and Pandas will be heavily used for this purpose. The approach will use jackknife resampling to produce MI estimates and the accompanying bias and variance.

II. Synthetic Data Generation and Experimentation: After the jackknife approach is developed, a series of controlled tests will be carried out on synthetic datasets. These datasets will be generated with varying degrees of dependency between variables, ranging from no dependency to heavy reliance, as well as varied amounts of noise to evaluate the estimator's resilience. The datasets will also differ in sample size to assess the method's performance in small and large data settings.

III. Comparison of Existing MI Estimation Techniques: In addition to the jackknife method, other MI estimating techniques will be used for comparison purposes. This will aid in determining the relative performance of the jackknife approach across various contexts and data kinds. The goal is to determine how successfully the jackknife strategy reduces bias, particularly in small sample sizes, and how computationally efficient it is compared to alternative methods.

IV. Application to Real-World Data (if applicable): In addition to synthetic data, the jackknife method will be aimed to be tested using a real-world dataset. This would be achievable if an appropriate dataset is available for the purpose. This will help to validate the method's viability in real-world applications.

V. Documentation and Final Report: Finally, the results and discoveries from the implementation and experimentation phases will be combined into a thorough final report. This report will summarize the study approach, present the findings, and analyze their implications, with a focus on prospective contributions to the field of mutual information estimation.

By the end of Postgraduate Project B, the entire implementation and evaluation of the jackknife approach will be concluded. The project's findings will help us understand the strengths and limitations of the jackknife method for estimating mutual information, as well as pave the way for future study towards better MI estimate methodologies.

References

Arya, N., 2023. *The Role of Resampling Techniques in Data Science - KDnuggets*. [Online]
Available at: https://www.kdnuggets.com/2023/02/role-resampling-techniques-data-science.html
[Accessed 5 October 2024].

Aznar, P., 2021. *What is Mutual Information_ _ Quantdare*. [Online]
Available at: https://quantdare.com/what-is-mutual-information/
[Accessed 1 October 2024].

Cover, T. M. & Thomas, J. A., 2006. *Elements of Information Theory.* 2nd ed. New Jersey: Wiley-Interscience.

Crosato, E. & Morris, R. G., 2023. *information_theory.* [Online]
Available at: https://www.emblaustralia.org/wp-content/uploads/2023/11/information_theory.pdf
[Accessed 2 October 2024].

Frost, J., 2024. *Introduction to Bootstrapping in Statistics with an Example - Statistics By Jim.* [Online]
Available at: https://statisticsbyjim.com/hypothesis-testing/bootstrapping/
[Accessed 5 October 2024].

Glen, S., 2019. *Resampling Methods_ Bootstrap vs jackknife - DataScienceCentral.com*. [Online]
Available at: https://www.datasciencecentral.com/resampling-methods-comparison/
[Accessed 5 October 2024].

Khan, S. et al., 2007. *neu_330006*. [Online]
Available at: https://repository.library.northeastern.edu/files/neu:330004/fulltext.pdf
[Accessed 5 October 2024].

Kraskov, A., Stögbauer, H. & Grassberger, P., 2004. Estimating mutual information. *Phys. Rev. E,* 69(6), p. 066138.

Moon, Y.-I., Rajagopalan, B. & Lall, U., 1995. Estimation of Mutual Information Using Kernel Density Estimators. *Physical review. E, Statistical physics, plasmas, fluids, and related interdisciplinary topics,* 52(3), pp. 2318-2321.

Shannon, C. E., 1948. A mathematical theory of communication. *The Bell System Technical Journal,* Volume 27, pp. 279-423.

Wikipedia, 2024. *Bootstrapping (statistics) - Wikipedia*. [Online]
Available at: https://en.wikipedia.org/wiki/Bootstrapping_(statistics)
[Accessed 5 October 2024].

Wikipedia, 2024. *Jackknife resampling - Wikipedia*. [Online]
Available at: https://en.wikipedia.org/wiki/Jackknife_resampling
[Accessed 5 October 2024].

Wikipedia, 2024. *Mutual information - Wikipedia*. [Online]
Available at: https://en.wikipedia.org/wiki/Mutual_information
[Accessed 1 October 2024].

Zeng, G., 2015. A Unified Definition of Mutual Information with Applications in Machine Learning. *Mathematical Problems in Engineering,* Volume 2015.

Zeng, X., Xia, Y. & Tong, H., 2018. *Jackknife approach to the estimation of mutual information _ PNAS*. [Online]
Available at: https://www.pnas.org/doi/10.1073/pnas.1715593115
[Accessed 1 October 2024].

Zhang, R., 2023. *Information Theory Series_ 1 — Entropy and Shannon Entropy _ by Renda Zhang _ Medium*. [Online]
Available at: https://rendazhang.medium.com/information-theory-series-1-entropy-and-shannon-entropy-a20a2101108e
[Accessed 2 October 2024].