2023 IEEE Symposium Series on Computational Intelligence (SSCI)
Mexico City, Mexico. December 5-8, 2023

Comparative Analyzes of Human and Machine Randomness: Insights into Decision-Making Models

Sofia Marshallowitz*, Edison Pignaton de Freitas*

*Federal University of Rio Grande do Sul, Brazil
stapuzzo@ufrgs.br, edison.pignaton@inf.ufrgs.br

***Abstract*—Human decision theory focuses on the reasoning behind the choices an individual makes. Human decision modelling is developed through mental models and can be modelled in different ways, such as fuzzy logic, deductive logic and probabilistic logic. On the other hand, machine learning techniques use a variety of statistical, probabilistic, and optimization methods to learn and detect useful patterns. In this context, this study investigates the complexities of human and machine randomness, utilizing two distinct datasets: one representing the perceived randomness of humans through the selection of nine numbers and the other encapsulating algorithmically generated random numbers from machines. The comparison of these datasets aims to understand the similarities and divergences between human (brain) randomness and machine randomness, primarily through the lens of fairness, neurocomputational, and decision-making simulations.**

***Index Terms*—Neuromathematics, Randomness, Ensemble learning, Human–machine comparison, Neurocomputing**

I. Introduction

The ability to make human decisions is not perfect, and even with the same information, different decisions can be reached. This is called variability [1], and it occurs both between a group of experts in an area and between decisions made by a single expert. Human decision theory is concerned with the reasoning behind an individual's choices, whether it is a simple choice between whether to sweeten the coffee or not, or a deeper choice, such as whether to marry or not, or, considering the story of Ulysses, whether to return to Ithaca by the simplest and safest route or to take the risk to hear the sirens' song before returning, and submit to a plan that just one slip could be fatal, which is as complex to explain as deep machine learning models. The modelling of decision-making is developed through mental models and, as with machine learning algorithms, there are distinct approaches such as fuzzy-logic, deductive logic and probabilistic for its modelling, each with effectiveness, although the approach can be common in both scenarios.

However, these choices can find a ”noise”: the random. Randomness, in the common understanding, occurs when outcomes transpire haphazardly, unpredictably, or by sheer chance. Though these three concepts are distinct, they all bear a close relationship to probability. Notably, probability comes in various forms: subjective probabilities (or 'degrees of belief'), evidential probabilities, and objective chances, among others. One might explore the ties between randomness and any of these types of probability. In this context, the focus of this work is on the possible connections between randomness and chance, also referred to as physical probability. The typical usage of the word 'random' tends to be more or less synonymous with 'chancy' daily. [2]

Decision-making is a cognitive process of central importance. Increasing evidence indicates that behavioral variability is crucial in the way humans balance the trade-off between exploration and exploitation. In making these decisions, a slight amount of variability can assist humans in resisting the urge to exploit known rewards, instead prompting them to explore other options at random [3]. Determining when to continue exploring or when to halt and utilize what is available is a vital component in many decisions, grasping what governs random exploration can enhance decision-making across various facets of human life.

Complementary, human randomness perception is commonly described as biased [4] once it occurs because when humans create random sequences, they often consistently under-represent or over-represent specific subsequences, compared to what would be anticipated from a genuinely unbiased random process. Randomness is indifferent to history and humans have one, constantly flirting with their subjectivity. The search for ”true” randomization is something old, since in Athens, around 300 BC, right at the beginning of what is called democracy today, elections did not involve votes in the way popularized today, but rather there was the use of the kleroterion, a device of randomization consisting of a carved stone plate with rows of slits and an attached tube. [5] This study used a sequence of numbers created by a Random Number Generation (RNG) algorithm. For a number in a sequence or distribution to be truly random, it must be independent. The independence of numbers means there is no correlation between successive numbers. In addition, these numbers should occur in the distribution with approximately the same frequency. [6]

However, one point that approximates the absence of true randomness between machines and humans is that computers currently also lack true randomization and RNG actually deals with pseudorandomization, which is intrinsic to the deterministic nature of computers. [7]

While software-generated random sequences are not genuinely random, viable alternatives can be found in fast entropy sources like quantum systems or classically chaotic systems, as long as they can create high-quality random sequences at a sufficient speed. The unearthing of spontaneous chaos in semiconductor superlattices at room temperature has given rise to a valuable option in the field of nanotechnology, emphasizing that the machines and computational structures used nowadays do not have the idealized random process. [8]

In this context, the goal of this study is to perform an investigation on the complexities of human and machine randomness. This study uses two distinct datasets: one representing the perceived randomness of humans through the selection of nine numbers and the other encapsulating algorithmically generated random numbers from machines. Aiming to understand the similarities and divergences between human (brain) randomness and machine randomness, these datasets are compared with neurocomputational, and decision-making simulations.

II. Related Work

Within the field of human decision-making, there is strong behavioral and physiological evidence that the brain represents probability distributions and performs probabilistic inference [9]. Some techniques place them as complementary methods and experimental evidence that supports the notion that human behavior is highly consistent with Bayesian probabilistic inference in the sensory as well as motor and cognitive domain (according to the Bayesian definition, probabilities are personal beliefs). On the other hand, from the frequentist point-of-view, the reality of the world is that it will or will not rain tomorrow: it makes no sense to a human that "it will rain 60%." So, continues the argument, from a realist point of view, single-case probabilities are meaningless. If one wants to make sense of probabilities, the only way to do so is to treat probabilities as frequencies: the subjective, Bayesian understanding of probabilities is hopelessly subjective and, as such, should not be considered. [10].

The fuzzy strategy imports topics that flirt with philosophy and unorthodox lines of understanding machine decision effectiveness [11], which discusses that if the problem domain is such that human experts cannot achieve 100% performance. A computer expert system in this domain should not be expected to do so; that is: if human experts are allowed to make mistakes, then it should be allowed a computer expert system to do so, so that the reasoning is equivalent, even in the misunderstandings.

More metaphysical issues are those like the comparison between Bayesian and frequentist means [12]. Bayesian statistics, predating its frequentist counterpart, encapsulates all necessary information for inference within observed data, excluding unobserved variables. Historically sidelined due to its limitations in solving cases without known conjugate priors, Bayesian methods are witnessing a resurgence. This revival is fueled by advancements in IT and novel mathematical methodologies, coinciding with the rise of machine learning, where the confluence of statistics and computation propels statistical algorithms' relevance within this domain.

III. Problem Formulation Section

This section details the elements of the methodology used in the problem formulation in the proposed study, i.e., the datasets and the proposed modelling.

A. Datasets

Two datasets were used, named $human\ decision\ dataset$ and $machine - generated\ dataset$. Both consist of 20 columns with 50 rows made up of numbers from 1 to 10 on each register.

The human decision dataset is available on Kaggle [12] under a Creative Commons license and it was carried out through a survey of 20 different and anonymous individuals, worldwide, therefore without the description of age, gender, or other qualifications. The machine-generated dataset consists of an algorithm in the software R based on $runif()$, a function to generate random values from a uniform distribution in R.

B. Modelling

For preliminary analysis, techniques were used to visualize normality distribution, followed by probability density function (PDF) and statistical analyzes, which in the context of this study will be called preliminary analyzes. To reinforce the exploration, it was also used the XGBooster [13], RandomForest [14], Support Vector Regression [15], and Ridge Regression [16], evaluated by the loss check using Mean Absolution Error.

IV. Used Methods

This section details the elements of the methodology used to handle the modelling in addition to statistical analyzes.

1) Normality Distribution: In examining a process where numbers are randomly selected from the range of 1 to 10 for a total of nine trials, the underlying statistical distribution that governs this process can be identified as a uniform distribution. This is predicated on the assumption that each individual number within the specified range is equally likely to be chosen in any given trial, leading to a probability of $\frac{1}{10}$for each number.

An interesting note, however, is that the Central Limit Theorem may come into play if one were to repeatedly take samples of size nine from this uniform distribution and calculate their means. In such a scenario, as the number of sample means increases, their distribution would tend to a normal distribution, irrespective of the original uniform distribution from which the samples were drawn.

In conclusion, the process of randomly selecting a number from 1 to 10 for nine trials aligns with the uniform distribution, reflecting the equal probability for each potential outcome within the defined range.

2) Probability Density Functions: In continuous random variables, density is articulated via a probability density function (PDF), indicating the probability of a variable taking certain values. Despite the discrete appearance of selecting numbers 1 to 10, examining density here is enlightening.

In our uniform distribution, the probability mass function (PMF) remains constant, visualized as equal-height bars in a histogram, reflecting distribution uniformity. Transitioning this to a continuous context, we encounter the constant-density PDF of a continuous uniform distribution.This density, constant across the interval, is akin to a ”discrete density” in our scenario, represented by the PMF.

In summary, density, embodied in continuous distributions' PDF and discrete ones' PMF, is crucial for comprehending random phenomena, revealing outcome likelihoods and distribution structure. For our process, constant density emphasizes uniformity, while a shift to bell-shaped density would indicate intriguing large-sample dynamics.

3) XGBooster: XGBoost, an advanced ensemble learning technique, is pivotal in predictive modeling, leveraging sequential construction of weak learners, typically decision trees, and integrating L1 (Lasso) and L2 (Ridge) regularization to curb overfitting by penalizing model complexity.

In contexts of inherent randomness in datasets, traditional linear models may falter due to the absence of discernible patterns. However, XGBoost's ensemble approach and gradient boosting capabilities excel at uncovering latent structures within such data. The method's progressive learning from weak learners refines predictions, while its regularization avoids noise adherence, favoring genuine pattern recognition and overfitting prevention.

4) Random Forest: Random Forest, an ensemble method used for classification and regression, builds multiple decision trees during training and averages their predictions for regression tasks, inherently embracing randomness via its mechanisms.

This randomness is twofold: first, through Bootstrap Aggregating (Bagging), each tree trains on a unique random data subset, ensuring diverse tree formation. Second, random Feature Selection at each node during tree construction brings additional variability, reducing overfitting. In scenarios involving random numbers, Random Forest's stochastic nature may offer enhanced compatibility with the data's randomness, facilitating effective modeling. Its capacity for capturing non-linear complexities in seemingly chaotic random number data lies in aggregating diverse tree predictions, each based on different data and feature subsets, thereby achieving a comprehensive, robust data representation.

5) Support Vector Regression: Support Vector Regression (SVR), a variant of Support Vector Machine (SVM), addresses regression problems by identifying an optimal hyperplane within a specified error margin, using a distinctive feature: the kernel function. This function enables non-linear mapping of data, enhancing SVR's adaptability to complex patterns, crucial for randomness-inherent data.

SVR's resilience to randomness and noise stems from its emphasis on support vectors—data points defining the hyperplane—minimizing sensitivity to outliers. In random number datasets, SVR's versatility shines; the kernel (linear, polynomial, RBF) can be tailored to the data's randomness nature, accurately modeling complex, non-linear patterns.

Additionally, SVR's hyperparameters control error tolerance and penalties, crucial for balancing fit and overfit prevention in random data modeling. This balance is vital; overly stringent models might overlook inherent randomness patterns, while excessively lenient ones might overfit to random noise.

6) Ridge Regression: Ridge Regression, known for L2 regularization, addresses multicollinearity and overfitting in high-dimensional data by adding a penalty proportional to the squared L2 norm of its coefficients to the ordinary least squares (OLS) loss function. This addition, regulated by a hyperparameter λ, biases the model towards smaller coefficients, balancing data fit and coefficient magnitude.

In the context of random numbers, Ridge Regression is advantageous. Randomness can introduce noise, instability, and potential collinearity, all mitigated by Ridge Regression's regularization. The L2 focus evenly distributes predictor influence, fostering a stable model that captures randomness subtleties. The λ hyperparameter fine-tunes sensitivity to randomness, discerning genuine patterns amidst noise—a balance achieved through methods like cross-validation.

A. Mean Absolute Error

The Mean Absolute Error (MAE) is a straightforward error metric, averaging the absolute discrepancies between actual and predicted data points. Unlike the Mean Squared Error (MSE) with its quadratic penalty for errors, MAE linearly penalizes deviations, providing an intuitive gauge of average error magnitude, thereby simplifying model performance interpretation. Crucially, MAE's robustness against outliers due to its linear nature makes it apt for random number modeling, preventing disproportionate influence of extreme values inherent in MSE. By penalizing errors linearly, MAE ensures equitable treatment of all deviations, large or small, aligning well with contexts of unpredictability in random datasets. Moreover, if a model's optimization hinges on an absolute error framework, using MAE for evaluation maintains methodological coherence, streamlining both training and assessment processes by unifying objectives and performance metrics. This uniformity fosters transparency and enhanced reliability in modeling trajectories.

V. HUMAN DATASET RESULTS ON PRELIMINARY ANALYZES

In the provided context, the first three numbers were quite close to a normal distribution as can be seen in Fig 1.

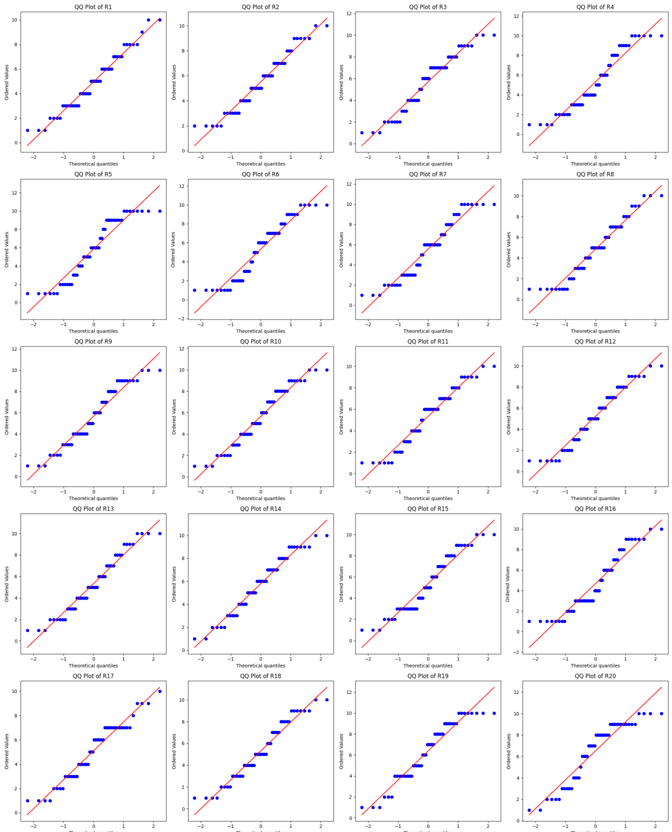

Fig. 1. Normality Analyzes

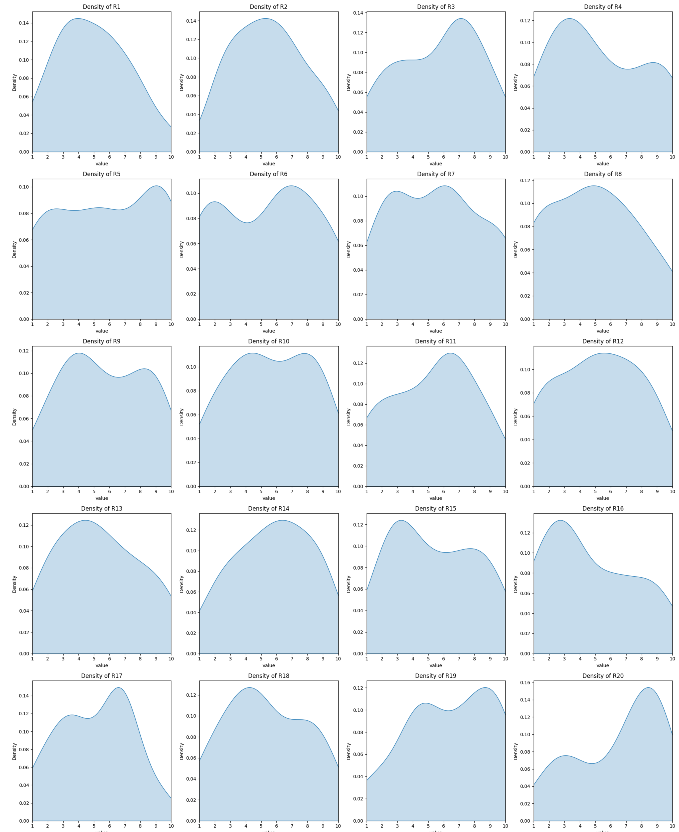

Fig. 2. Density Analyzes

However, upon further examination of the remaining numbers in the sequence, the characteristics changed, suggesting that these numbers were not following a normal distribution,

Then, the analysis involved the evaluation of densities is presented in Fig. 2. This analysis likely refers to the probability density functions associated with the different parts of the sequence. A non-normal distribution might display skewness (lack of symmetry) or kurtosis (presence of heavy tails or outliers) that would distinguish it from a normal distribution.

In this specific context, the highlighted that the individual seemed to exhibit a behavior of avoiding the repetition of the same number. The visualization likely demonstrated that the numbers were not repeated consecutively or perhaps even at regular intervals within the series, reinforcing the hypothesis that there was an intentional or systematic avoidance of number repetition.

This pattern could indicate a non-random process at play, as avoiding repeating numbers might suggest a conscious decision-making process rather than a truly random selection.

VI. Machine Generated Dataset Results on Preliminary Analyzes

A. Data Distribution

The analysis revealed that the second dataset was not so different in its behavior from the first one, with a distinction of similarity on ”central numbers”, as can be seen in Fig. 3.

In this case, initial and final numbers have a bit more of density than human choices. This finding might indicate that while there was a specific pattern or bias in the selection of central numbers, the underlying stochastic process generating these numbers was still predominantly random and followed the typical properties of a normal distribution. Therefore, the analysis of the second dataset revealed a pattern where the behavior was not entirely typical of a normal distribution, yet not so different that it could be classified as another type of distribution, as can be seen in Fig. 4. The nuanced characteristics, particularly the underrepresentation of central numbers, could provide valuable insights into the underlying mechanisms or biases at play, and may warrant further investigation using more specialized statistical techniques or models tailored to the specific context of the study.

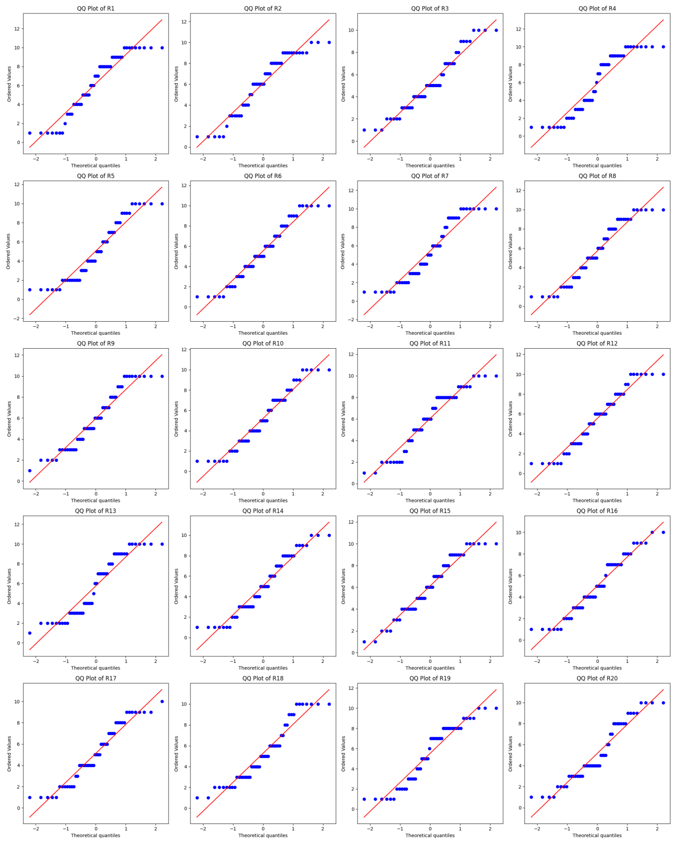

Fig. 3. Normality Analyzes

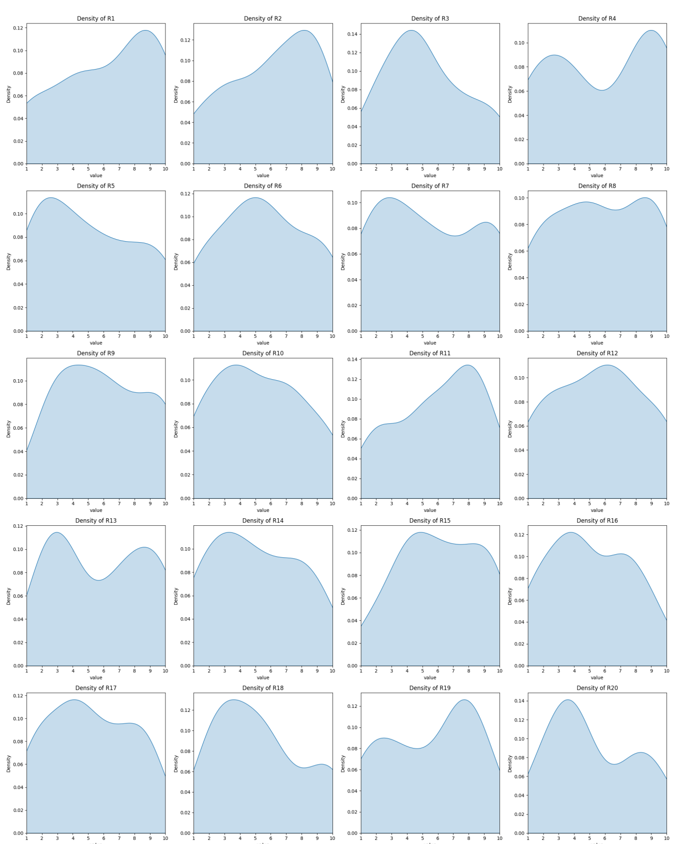

Fig. 4. Density Analyzes

VII. Statistical Analyzes

This section delves into a comprehensive statistical examination conducted to decipher inherent patterns and variances in the dataset concerning human choices and machine-generated responses. Utilizing an array of statistical methodologies, including descriptive statistics, hypothesis testing, and variance analysis, this segment aims to elucidate significant disparities or similarities between human and algorithmic behaviors in randomness generation.

Statistic	Human	Machine
Mean Range	4.66 - 6.50	5.02 - 6.20
Standard Deviation Range	2.29 - 3.22	2.64 - 3.35
T-test P-value	0.97	0.10
ANOVA F-value	1.26	81.57
ANOVA P-value	0.29	<0.0001
Sample Size Required	64	64

TABLE I
Comparative Summary of Statistical Analysis

A. *Descriptive Statistics*

The human participants' mean responses hovered between 4.66 (R16) and 6.50 (R20), whereas the machine-generated numbers presented a tighter cluster, with means ranging from 5.02 (R16) to 6.20 (R1). The human data showcased a broader range, hinting at possible cognitive biases or decision-making heuristics affecting number selection. The standard deviations, signifying response variability, were comparably distributed in both sets, albeit slightly higher in the human data, indicative of greater inconsistency and potential bias in human choices.

B. *T-test*

A T-test comparison between two randomly selected columns from the machine data resulted in a P-value of 0.09717545660261137, indicating no significant mean difference, suggesting a consistent distribution across selections. This consistency underscores the algorithm's efficacy in maintaining uniform randomness, just like humans.

C. *ANOVA*

However, the Analysis of Variance (ANOVA) painted a different picture. While the ANOVA on human data suggested no significant mean differences across groups (F=1.257941, P=0.28728), the machine data exhibited a stark contrast. With an F-value of 81.572661 and a P-value reaching statistical significance at 1.471565e-24, the test implied a substantial difference in at least one of the group means. This disparity between groups in the machine data underscores the algorithm's proficiency in simulating true randomness, surpassing the "human randomness".

VIII. Comparing MAE

The results of the metric to evaluate both human and machine prediction accuracy.

Despite the slight difference between the MAE values on the human dataset and the generated dataset, the first being less than the second, indicating a better accuracy and, therefore,

Model	Human MAE	Machine MAE
XGBoost	0.269	0.288
Random Forest	0.271	0.289
SVR	0.320	0.334
Ridge	0.260	0.291

TABLE II
COMPARISON OF MEAN ABSOLUTE ERROR FOR DIFFERENT MODELS

greater predictability (that is, the generated dataset presented a higher level of " randomness" or, at least, difficult to predict). The XGBoost model exhibited a commendable performance, characterized by a nominal disparity between human and machine error rates. This modest discrepancy underscores the model's capacity to align machine predictions closely with human judgment, thus making it an attractive option for applications requiring human-like reasoning. Similar to XGBoost, the Random Forest model demonstrated a consistent predictive performance, evidenced by a near-equal human and machine MAE. On the other hand, the SVR model manifested a higher MAE, suggesting a reduced alignment with human judgment. This divergence may necessitate a closer examination of model parameters, hyperparameter tuning, or exploration of alternative kernel methods to enhance predictive fidelity. Tuning was not used in this study. Finally, achieving the lowest human MAE among the examined models, Ridge Regression revealed a more pronounced gap between human and machine errors. The marginal disparities observed in the Mean Absolute Error (MAE) values between human and machine-based decisions across the deployed models bear significant implications for the interpretability and reliability of machine learning (ML) algorithms, particularly under conditions infused with inherent randomness. These discrepancies, albeit slight, serve as critical indicators of the models' robustness and their ability to emulate human cognitive processes within uncertain environments.

IX. CONCLUSION

This work proposed a comparative study on human-based and machine-based decision-making in randomness. This is an important subject as the advance of emerging technologies relying on machine-based decisions, it is important to understand in which circumstances decisions taken by computers can be similar, or not, to those taken by humans.

The study submitted two different datasets, one representing decisions taken by humans and another by a computer, for comparative analysis. The results demonstrate that sequential human decisions have some level of dependence on each other, thus canceling out a completely random choice. However, machines also do not operate in true randomness. The distance between human and computational randomness is not so long, thus allowing explorations in the field of neurocomputing and brain simulation that flirt with human reality and behavior.

As future work, it is possible to explore more complex datasets, representing decisions related to more complex situations, such as those found, for instance, in situations like driving a car or manipulating goods in a warehouse, typical tasks that are being delegated to computers (robots) in emerging applications. Thus understanding the differences between decisions taken by human operators and a computer is of crucial interest to the designers of such applications.

REFERENCES

[1] Jonathan M. Garibaldi. The need for fuzzy ai. *IEEE/CAA Journal of Automatica Sinica*, 6(3):610–622, 2019.

[2] Antony Eagle. Chance versus Randomness. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Spring 2021 edition, 2021.

[3] Christopher J. G. Meacham and Jonathan Weisberg. Representation theorems and the foundations of decision theory. *Australasian Journal of Philosophy*, 89(4):641–663, 2011.

[4] Jonathan M. Garibaldi. The need for fuzzy ai. *IEEE/CAA Journal of Automatica Sinica*, 6(3):610–622, 2019.

[5] Mehdi Pirooznia, Jack Y. Yang, Mary Qu Yang, and Youping Deng. A comparative study of different machine learning methods on microarray gene expression data. *BMC Genomics*, 9(1):S13, Mar 2008.

[6] Baihan Lin. Computational inference in cognitive science: Operational, societal and ethical considerations, 2022.

[7] Małgorzata Figurska, Maciej Stańczyk, and Kamil Kulesza. Humans cannot consciously generate random numbers sequences: Polemic study. *Medical Hypotheses*, 70(1):182–185, 2008.

[8] Amarjitsing Rajput, Ganesh Shevalkar, Krutika Pardeshi, and Prashant Pingale. Computational nanoscience and technology. *OpenNano*, 12:100147, 2023.

[9] Alexandre Pouget, Jeffrey M Beck, Wei Ji Ma, and Peter E Latham. Probabilistic brains: knowns and unknowns. *Natural Neuroscience*, 16(9):1170–1178, 2013.

[10] Fei Xu and Tamar Kushnir. Chapter one - the probable and the possible at 12 months: Intuitive reasoning about the uncertain future. In *Rational Constructivism in Cognitive Development*, volume 43 of *Advances in Child Development and Behavior*, pages 1–25. JAI, 2012.

[11] Francisco J. Samaniego. *A Comparison of the Bayesian and Frequentist Approaches to Estimation*. Springer, New York, 2010.

[12] Kaggle datasets: Can humans really be random? https://www.kaggle.com/datasets/passwordclassified/can-humans-really-be-random. Accessed: 2023-08-01.

[13] Tianqi Chen and Carlos Guestrin. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '16, page 785–794, New York, NY, USA, 2016. Association for Computing Machinery.

[14] Leo Breiman. Random forests. *Machine Learning*, 45(1):5–32, Oct 2001.

[15] Corinna Cortes and Vladimir Vapnik. Support-vector networks. *Machine Learning*, 20(3):273–297, Sep 1995.

[16] Donald E. Hilt, Donald W. Seegrist, United States. Forest Service., and Pa.) Northeastern Forest Experiment Station (Radnor. *Ridge, a computer program for calculating ridge regression estimates*, volume no.236. Upper Darby, Pa, Dept. of Agriculture, Forest Service, Northeastern Forest Experiment Station, 1977. https://www.biodiversitylibrary.org/bibliography/68934.

ACKNOWLEDGEMENT

The authors gratefully acknowledge the financial contribution of an anonymous donor who prefers to be identified only as S.W for registration fees for this article. The first author, in particular, also thanks S.W for the valuable conversations and the users who got involved in the discussion on the Kaggle website about the dataset used here.