Nonparametric AI Innovations for Social Science

Sthitadhi Das
Assistant Professor, Department of Mathematics, Brainware University, India

Abstract: Social science research increasingly relies on sophisticated analytical tools to uncover complex, nonlinear patterns in large-scale and heterogeneous datasets. This review explores the intersection of artificial intelligence (AI) and nonparametric statistical estimation within the social sciences. Emphasis is placed on machine learning (ML) and deep learning (DL) methods that extend classical statistical techniques, enabling researchers to analyze social phenomena without imposing rigid parametric assumptions. Methodological innovations—such as deep kernel learning, reinforcement learning for model tuning, and neural additive models—are examined for their application in areas like public opinion modeling, income inequality, educational outcomes, and behavioral prediction. We also discuss theoretical developments, including consistency, convergence, and generalization, that support the integration of AI into nonparametric frameworks. Challenges such as interpretability versus accuracy, computational costs, and ethical considerations are addressed. We conclude by outlining future research directions, including hybrid modeling, fairness-aware inference, and privacy-preserving analytics in social data contexts.
Keywords: Artificial Intelligence (AI); Social Science; Nonparametric Estimation; Machine Learning (ML); Deep Learning (DL); Behavioral Prediction; Kernel Methods; Neural Additive Models; Interpretability; Ethics in AI.

Introduction

The use of data-driven methodologies in the social sciences has expanded significantly over the past two decades, driven by both the availability of complex data sources—such as social media, administrative records, and longitudinal surveys—and the development of powerful computational tools. Traditional parametric models, while foundational, often fall short in capturing nonlinearities, heterogeneity, and interactions common in social phenomena.

Nonparametric estimation, which avoids strict functional assumptions, offers a natural framework for flexible modeling. When paired with artificial intelligence (AI) methods, especially machine learning (ML) and deep learning (DL), these techniques provide a robust toolkit for understanding social behavior, policy effects, and demographic trends.

This review surveys recent methodological advances at the intersection of nonparametric statistics and AI in the context of social science applications. Topics such as public health, economics, sociology, and education research serve as case studies to illustrate the practical utility of these methods. We highlight the importance of balancing predictive performance with interpretability and discuss how ethical concerns, such as algorithmic bias and data privacy, are increasingly shaping methodological choices.

The integration of AI into social science research has led to a paradigm shift, enabling researchers to analyze complex datasets with greater precision and efficiency. AI techniques, including natural language processing and neural networks, facilitate the extraction of meaningful patterns from unstructured data, such as text and images, thereby enriching the analytical toolkit available to social scientists.

Furthermore, the application of nonparametric methods in AI models enhances their adaptability and robustness, particularly in scenarios where traditional parametric assumptions do not hold. By leveraging these methods, researchers can develop more accurate models that better reflect the underlying complexities of social systems, leading to more informed policy decisions and interventions.

Literature Review: Classical nonparametric techniques such as kernel density estimation [9], local regression [5], and spline smoothing [11] have long been used in social statistics to study income distributions, vote shares, or fertility patterns. However, as datasets have grown in complexity, machine learning has emerged as a complementary tool.

Neural networks, initially criticized for their black-box nature, are now used in modeling latent psychological traits and predicting recidivism risk [10]. Random forests and gradient boosting machines have gained popularity for imputing missing survey data and predicting education attainment [1]. Bayesian nonparametric provide interpretable uncertainty quantification in models of political opinion or migration [7].

AI-driven advancements—like deep kernel learning [12] or reinforcement learning for survey design [14] —enable flexible, automated, and scalable analyses. Social scientists are also adopting interpretable ML models such as generalized additive models and SHAP-based explanations to retain transparency [8].

Our contribution is to synthesize these developments through a social science lens, linking methodological innovations with pressing questions in inequality, education, labor, and behavioral science.

Recent surveys have categorized the intersection of AI and social science into two primary directions: AI for social science, where AI tools enhance various stages of social science research, and social science of AI, which examines AI agents as social entities with human-like cognitive and linguistic capabilities [13]. This bifurcation underscores the dual role of AI in both advancing research methodologies and serving as a subject of social inquiry.

Furthermore, the application of nonparametric methods in AI models enhances their adaptability and robustness, particularly in scenarios where traditional parametric assumptions do not hold. By leveraging these methods, researchers can develop more accurate models that better reflect the underlying complexities of social systems, leading to more informed policy decisions and interventions.

Nonparametric Estimation Techniques

Kernel Density Estimation

Kernel Density Estimation (KDE) is a widely used nonparametric method for estimating the probability density function of a random variable. Given a sample $X_1, \ldots, X_n$, the KDE at a point x is defined as:

$$\hat{f}(x) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right)$$

where K is the kernel function (commonly Gaussian), and h is the bandwidth parameter that controls the smoothness of the estimate. The choice of h is crucial; too small a value leads to overfitting, while too large a value results in over-smoothing. A commonly used rule of thumb for selecting h is Silverman's method:

$$h = \left(\frac{4\hat{\sigma}^5}{3n}\right)^{1/5}$$

where $\hat{\sigma}$ is the sample standard deviation and n is the sample size. KDE is particularly useful in visualizing multimodal distributions and identifying underlying patterns in data.

Nonparametric Regression

Nonparametric regression aims to estimate the relationship between a dependent variable Y and an independent variable X without assuming a specific parametric form for the regression function. The Nadaraya-Watson kernel estimator is a popular method:

$$\hat{m}(x) = \frac{\sum_{i=1}^{n} K\left(\frac{x-X_i}{h}\right) Y_i}{\sum_{i=1}^{n} K\left(\frac{x-X_i}{h}\right)}$$

where K is the kernel function and h is the bandwidth parameter. This estimator provides a smooth estimate of the regression function and is particularly effective in capturing nonlinear relationships in data.

Spline Estimation

Spline estimation involves fitting piecewise polynomial functions to data, ensuring smoothness at the boundaries of the pieces. Cubic splines are commonly used, where the function is piecewise cubic and the first and second derivatives are continuous. The estimation involves minimizing the penalized residual sum of squares:

$$\sum_{i=1}^{n}\left(Y_i - s(X_i)\right)^2 + \lambda \int \left(s''(x)\right)^2 dx$$

where $s(x)$ is the spline function and λ is the smoothing parameter that controls the trade-off between fit and smoothness. Spline estimation is widely used in modeling nonlinear trends in data, such as demographic changes or policy effects.

k-Nearest Neighbors (k-NN)

The k-Nearest Neighbors method is a nonparametric technique used for both classification and regression. In regression, the estimator at a point x is the average of the k nearest neighbors:

$$\widehat{m}(x) = \frac{1}{k} \sum_{i \in \mathcal{N}_k(x)} Y_i$$

where $\mathcal{N}_k(x)$ denotes the set of indices corresponding to the k nearest neighbors of x. The choice of k is important; too small a value can lead to overfitting, while too large a value can oversmooth the estimate. k-NN is effective in capturing local patterns in data and is widely used in various applications, including social behaviour prediction.

Rank-based Methods

Rank-based methods are particularly useful for analyzing ordinal data, such as survey responses. These methods involve ranking the data and analyzing the ranks rather than the raw values. Examples include the Mann–Whitney U test and the Kruskal-Wallis test, which are nonparametric tests used to compare differences between two or more groups. These methods do not assume normality and are robust to outliers, making them suitable for social science applications where data may not meet parametric assumptions.

Empirical Likelihood and Bootstrap Methods

Empirical likelihood (EL) is a nonparametric method for statistical inference that does not assume a specific parametric model. It involves constructing a likelihood function based on the empirical distribution of the data and maximizing it subject to certain constraints. EL methods are particularly useful for constructing confidence intervals and conducting hypothesis tests without relying on parametric assumptions.

Bootstrap methods involve resampling the data with replacement to estimate the sampling distribution of a statistic. This approach allows for the estimation of standard errors, confidence intervals, and hypothesis tests without relying on parametric assumptions. Bootstrap methods are widely used in social science research for uncertainty quantification and model validation.

Numerical Examples

Income Distribution Modeling

Consider simulating a bimodal income distribution:

$$X \sim 0.6 \cdot \mathcal{N}(20, 5^2) + 0.4 \cdot \mathcal{N}(60, 10^2)$$

Applying Kernel Density Estimation to this data reveals the presence of two distinct income groups, highlighting income inequality and class segmentation. This approach provides a clear visualization of the distribution and can inform policy decisions aimed at addressing disparities.

Educational Achievement vs Income

Simulate data where educational achievement Y is related to household income X:

$$Y = \log(1 + X) + \varepsilon, \quad \varepsilon \sim \mathcal{N}(0,1)$$

Using nonparametric regression to estimate the relationship between Y and X reveals a nonlinear trend, indicating that the effect of income on educational achievement is not constant but varies across different income levels. This insight can inform targeted interventions in education policy.

Software Implementations

Several software packages provide implementations of nonparametric estimation techniques:

- R: The np package offers functions for kernel density estimation and nonparametric regression.
- Python: The scikit-learn library provides implementations of k-NN and other nonparametric methods.

- MATLAB: The fit function supports spline fitting and other nonparametric techniques.

These tools facilitate the application of nonparametric methods in social science research, enabling researchers to analyze complex data without relying on restrictive parametric assumptions.

Theoretical Foundations of AI-Enhanced Nonparametric Methods in Social Science

The convergence of artificial intelligence (AI) and classical nonparametric inference has led to the development of novel theoretical frameworks that are particularly suited to the challenges of social science data—namely, high dimensionality, heterogeneity, and weak prior structure. Traditional nonparametric methods rely on minimal assumptions and offer flexible estimation tools. However, their performance often degrades in high-dimensional settings due to the curse of dimensionality. AI models, particularly those based on deep learning, mitigate this through data-adaptive representation learning, and recent work has provided a growing foundation for understanding their statistical properties.

Neural Networks as Nonparametric Estimators

Neural networks are now commonly viewed through the lens of nonparametric function estimation. Their ability to approximate any Borel-measurable function—under sufficient width or depth—positions them as powerful alternatives to classical kernel or spline methods. In social science applications, such as modelling voter preferences, ideological positioning, or inequality dynamics, neural networks capture nonlinear patterns that are often missed by standard techniques.

From a theoretical standpoint, consistency and convergence of neural networks depend on the number of parameters relative to the sample size, the architecture (depth, width), and the regularization mechanism (explicit or implicit). Recent studies [2,4,15] have shown that under mild conditions, multilayer networks trained via gradient descent exhibit consistency, and generalization bounds can be established using tools from empirical process theory, covering numbers, and norm-based capacity control.

Benign Overfitting and Double Descent

One of the most profound shifts in modern statistical thinking comes from the study of overparameterized models. Classical theory warns against models that interpolate the training data, expecting poor generalization due to overfitting.

However, recent theoretical breakthroughs have revealed that under specific conditions particularly with structured noise, proper initialization, and inductive bias of optimization algorithms—overparameterized models can generalize well. This phenomenon, known as benign overfitting, has implications for survey-based modelling, panel data analysis, and behavioural forecasting in the social sciences.
The so-called double descent curve, where risk initially increases with model complexity but later decreases beyond the interpolation threshold, provides a new lens for understanding regularization in deep learning. These insights motivate researchers to reevaluate the classical bias–variance trade-off in light of algorithmic and structural regularization rather than explicit penalties.

Simulation Illustration: Nonlinear Attitudinal Modelling

To concretely demonstrate the theoretical insights on neural network behaviour in social science contexts, we conduct a detailed simulation study focusing on modelling nonlinear attitudes toward social policies. This example captures realistic phenomena such as how latent trust in institutions shapes support for policy measures in a non-additive, smooth but complex manner.

Simulation Setup

We generate synthetic data where the response variable Y represents an individual's degree of support for a given policy, modelled as a nonlinear function of a latent continuous variable X representing institutional trust or confidence. The data-generating process is specified as:
$Y = \sin(2\pi X) + \varepsilon$, where X follows a Uniform (0, 1) distribution and ε is Gaussian noise with mean 0 and variance 0.07.
This setting embodies realistic complexities: the sine function introduces periodic nonlinearities capturing fluctuations in support, while the noise term reflects variability in individual responses that social scientists often observe.

Neural Network Architecture and Training

We implement fully connected feedforward neural networks (multilayer perceptrons) with a single hidden layer using ReLU activation functions. To investigate the impact of model capacity, we vary the hidden layer width across three regimes: low capacity ($H = 5$), moderate capacity ($H = 20$), and high capacity or overparameterization ($H = 100$).
Networks are trained using stochastic gradient descent (SGD) to minimize mean squared error (MSE) loss over a training set of fixed size. To isolate the effect of network size on fitting and generalization, explicit regularization techniques (e.g.,

weight decay or dropout) are omitted. Early stopping based on validation loss is employed to prevent overfitting.

Results and Interpretation

The trained models exhibit markedly different behaviours depending on the network width:

- For H = 5, the model underfits the data, failing to capture the oscillatory pattern in Y. The fitted function is overly smooth and approximates only a linear trend, resulting in high bias.
- Increasing the width to H = 20 enables the network to flexibly approximate the nonlinear sine curve. Both amplitude and phase are captured, leading to low bias and acceptable variance.
- At H = 100, the model enters the overparameterized regime, nearly interpolating the training data including noise. Despite this, test performance remains strong due to implicit regularization a phenomenon known as benign overfitting.

Theoretical Implications

This simulation illustrates modern theoretical results on generalization behaviour in deep learning. Recent Researches by [3], [4] and [6] have formalized the conditions under which overparameterized networks trained via gradient descent achieve consistency and low risk.

Key mechanisms include the implicit bias of gradient descent and norm-based control, which guide solutions toward smoother, lower-complexity functions. These findings confirm that high-capacity AI models can be both flexible and statistically reliable for behavioural research.

Thoughts

The theoretical foundations of AI-enhanced nonparametric methods offer a principled basis for analyzing social data. As data grows in complexity and scale, these methods will become more essential. Future directions include better uncertainty quantification, improved regularization strategies, and integration of these methods into user-friendly tools for the social sciences.

Methodological Innovations in Social Science Estimation and Inference

Recent advances in statistical methodology have introduced a powerful synergy between nonparametric inference and artificial intelligence, yielding tools that are both flexible and interpretable features especially crucial in social science applications where transparency and policy accountability are central. This section outlines several cutting-edge innovations that extend classical methods using neural networks, optimization techniques, and principled inference frameworks.

A notable development is the rise of deep kernel learning (DKL), which fuses neural network feature extraction with kernel-based regression or classification. Unlike standard kernel methods that rely on pre-specified similarity measures, DKL architectures allow the kernel to be learned from data through neural embeddings, typically optimized within a Gaussian Process (GP) or kernel ridge regression (KRR) framework. In reproducing kernel Hilbert space (RKHS)-based settings, the final kernel function becomes a composite of a neural transformation and a base kernel such as the radial basis function (RBF). This approach is particularly effective for structured covariates common in social datasets, such as census or longitudinal survey data.

Equally transformative is the use of reinforcement learning (RL) and Bayesian optimization (BO) for automating hyperparameter tuning. Tasks like bandwidth selection in kernel density estimation or penalty term calibration in spline regression are framed as sequential decision problems, where performance metrics (e.g., cross-validated error) guide updates via a reward signal. BO further employs Gaussian process surrogates to model the loss surface and propose efficient sampling points via acquisition functions like expected improvement. These automated strategies enhance reproducibility and performance, especially in large-scale studies with many preprocessing steps.

In the realm of inference, empirical likelihood (EL) has been extended to complex, nonparametric models including deep learners. Classical EL constructs likelihood ratios without fully specifying the distribution of the data, relying instead on moment constraints. Modern extensions replace these constraints with those implied by fitted black-box models—yielding nonparametric confidence regions for quantities like conditional means or treatment effects. Coupled with resampling methods such as the bootstrap or subsampling, these approaches permit valid post-estimation inference even when the estimator has no closed-form distribution.

Another key innovation is the use of neural additive models (NAMs), which build on the interpretability of generalized additive models (GAMs) while harnessing the representational power of neural networks. In NAMs, each covariate contributes through its own subnetwork, and the final prediction is obtained by summing these outputs. This additive structure not only ensures transparency in feature contribution critical for fairness and policy justification but also allows for visualization of learned effects through smooth functional plots. Regularization strategies such as group sparsity or monotonicity constraints can be incorporated to reflect domain knowledge or enforce fairness.

Case Studies in Methodological Innovation

Study 1: Deep Kernel Learning in Predicting Educational Attainment

We simulate data to represent educational performance influenced by nonlinear cognitive and social covariates.

Two models are fit for comparison:

- Classical kernel regression using a fixed RBF kernel.
- Deep kernel regression using a neural network feature extractor and a Gaussian process output layer.

While the classical method relies on Euclidean distances in the input space, the deep kernel model adapts to complex manifolds learned from data. Empirically, the latter shows lower prediction error and better generalization, reflecting its ability to capture latent interactions that affect academic outcomes.

Study 2: Neural Additive Models in Welfare Policy Evaluation

To mimic welfare allocation mechanisms based on household characteristics, we simulate a score function where the covariates may correspond to income, household size, and age, respectively.

A NAM is trained with separate subnetworks per feature, each constrained to be smooth and interpretable. Visual inspection of each function component reveals how marginal changes in inputs affect the welfare score—facilitating transparent allocation rules and aiding fairness audits.

Study 3: Empirical Likelihood for Inference in Deep Models

Using the setup from Study 1, we focus on a point estimate of the conditional expectation $E[Y|X = x_0]$ at $x_0 = (0, 0)$. Standard Gaussian process regression provides a posterior mean and variance, but for frequentist inference, we apply empirical likelihood based on residuals from the fitted model. This produces a nonparametric confidence interval that remains valid without assuming normality or linearity. Simulation experiments confirm that the coverage rates match theoretical expectations, demonstrating the potential for principled inference in black-box systems.

Overall, these methodological advances represent a paradigm shift in social science estimation. They offer ways to capture the richness of real-world data while preserving the interpretability and inferential rigor needed for transparent and credible scientific communication.

AI-Enabled Applications in Social Science Research

The fusion of artificial intelligence with nonparametric estimation techniques has significantly expanded the analytical capabilities of social scientists. By moving beyond rigid parametric assumptions, these hybrid methods enable more flexible, data-driven modelling of complex societal phenomena. In this section, we highlight three domains public health, economic inequality, and environmental justice where AI-enabled nonparametric estimation has proven especially impactful.

In public health, risk prediction and policy evaluation have benefited from the adoption of deep survival models such as DeepSurv. These architectures extend classical Cox proportional hazards models by allowing nonlinear covariate interactions and accommodating time-varying effects. As a result, they can more accurately estimate individualized health risks across diverse demographic groups defined by age, gender, income, or geographic region. These models have been used in pandemic-related research to assess vulnerability among high-risk populations, with uncertainty quantified using ensemble methods or Bayesian techniques. Such advances are particularly valuable in guiding public resource allocation and policy planning under uncertain conditions.

In the study of economic inequality and mobility, machine learning methods like quantile regression forests and generalized random forests have enabled fine-grained analysis of policy interventions across the income spectrum. Unlike linear regression or standard econometric models, these tree-based approaches adapt to data heterogeneity and nonlinear treatment effects, making them well-suited for estimating conditional quantiles or heterogeneous treatment effects. They have been employed to study long-run trends in wage inequality, intergenerational mobility, and the distributive consequences of taxation and social welfare programs. Their capacity to model distributional dynamics rather than just mean effects aligns with the growing demand for equity-aware analytics in economic research.

Environmental justice research has also seen a methodological transformation through the use of nonparametric models enriched by deep learning. Gaussian process regression, when combined with deep kernel learning, provides a flexible tool for spatiotemporal prediction of environmental exposures such as air pollution or toxic waste distribution. These models have been instrumental in mapping fine-grained pollutant concentrations—e.g., $PM_{2.5}$ or ozone levels—across regions disproportionately inhabited by marginalized communities. The probabilistic nature of these models allows not only point estimation but also principled uncertainty quantification, thereby supporting policy decisions that require a high degree of evidential robustness.

Taken together, these applications demonstrate the transformative potential of AI-augmented nonparametric estimation in addressing socially significant, high-dimensional problems. They show that with appropriate methodological integration, it is possible to retain statistical rigor, ensure model transparency, and generate actionable insights across critical areas of public concern. As these techniques continue to evolve, their adoption promises to deepen empirical understanding and inform more effective, equitable social policies.

Conclusion

The integration of artificial intelligence with nonparametric estimation is fundamentally transforming social science research. By harnessing the flexibility of nonparametric techniques alongside the scalability, adaptiveness, and representational power of AI, researchers are now able to model high-dimensional, nonlinear, and heterogeneous data structures that were previously intractable. This development has opened new avenues for empirical inquiry in domains such as public health, income inequality, environmental justice, and behavioural modelling.

Throughout this review, we have shown how AI-driven approaches including deep neural networks, kernel-based methods, and ensemble learning can augment classical estimation techniques by offering robust, data-adaptive alternatives. These innovations are supported by advancements in stochastic optimization, GPU-based computation, and automatic differentiation, which make it feasible to implement such methods on large-scale social science datasets.

Despite their promise, these methodologies introduce important challenges that must be addressed to ensure their broader applicability. Among these are concerns about model interpretability, computational efficiency, and theoretical guarantees related to consistency, convergence, and valid inference. Moreover, as these tools become increasingly influential in shaping public policy and social interventions, issues of transparency, fairness, and ethical accountability grow ever more critical.

Looking forward, it is essential that AI models for nonparametric inference in social sciences evolve to meet both statistical and societal demands. This includes the development of theoretically grounded estimation procedures that maintain the rigorous standards of classical statistics, while also incorporating flexible architectures suited to modern data. Furthermore, there is a growing need for scalable and reproducible workflows that can be shared across research communities, alongside the expansion of explainable AI frameworks that demystify the behaviour of complex models in socially meaningful terms. Ethical considerations, including fairness-aware inference and privacy-preserving analytics, must also be embedded in the design of new methodologies.

In conclusion, the fusion of artificial intelligence and nonparametric statistics holds the potential to create a new class of tools that are simultaneously powerful, interpretable, and socially responsible. As this interdisciplinary field continues to evolve, it is poised to play a pivotal role in driving empirical discovery and policy innovation across the social sciences.

Future Directions

The intersection of artificial intelligence and nonparametric estimation presents a rich landscape for future research in social sciences. Several promising avenues stand out:

- Developing theoretically grounded AI methods with provable consistency and robust uncertainty quantification.
- Designing scalable and reproducible computational frameworks, leveraging distributed computing and federated learning.
- Enhancing interpretability and transparency through explainable AI (XAI), neural additive models, and visualization tools.
- Embedding fairness-aware algorithms, privacy-preserving analytics, and ethical governance into methodological development.
- Fostering interdisciplinary collaboration across statistics, computer science, social science, and ethics.

Pursuing these directions promises to advance both the science and societal application of AI-enhanced nonparametric estimation, enabling nuanced, data-driven insights into complex social systems.

Acknowledgments

The author appreciates the valuable feedback from colleagues and reviewers, which significantly improved the clarity and depth of this work.

Conflict of Interest

The authors declare that there is no conflict of interest with respect to the publication of this manuscript.

References

1. Athey, S., Tibshirani, J., & Wager, S. (2019). Generalized random forests. The Annals of Statistics, 47(2), 1148–1178.
2. Barron, A. R. (1993). Universal approximation bounds for superpositions of a sigmoidal function. IEEE Transactions on Information Theory, 39(3), 930–945.

3. Bartlett, P. L., Long, P. M., Lugosi, G., &Tsigler, A. (2020). Benign overfitting in linear regression. Proceedings of the National Academy of Sciences, 117(48), 30063–30070.
4. Belkin, M., Hsu, D., Ma, S., & Mandal, S. (2019). Reconciling modern machine learning and the bias–variance trade-off. Proceedings of the National Academy of Sciences, 116(32), 15849–15854.
5. Fan, J., & Gijbels, I. (1996). Local polynomial modelling and its applications. Journal of the Royal Statistical Society: Series B (Methodological), 58(2), 547–567.
6. Hastie, T., Montanari, A., Rosset, S., & Tibshirani, R. J. (2022). Surprises in high-dimensional ridgeless least squares interpolation. The Annals of Statistics, 50(2), 949–986.
7. Hjort, N. L., Holmes, C., Müller, P., & Walker, S. G. (2010). Bayesian nonparametrics. Cambridge University Press.
8. Molnar, C. (2020). Interpretable machine learning. Lulu.com.
9. Parzen, E. (1962). On estimation of a probability density function and mode. The Annals of Mathematical Statistics, 33(3), 1065–1076.
10. Rudin, C. (2019). Stop explaining black box models for high stakes decisions and use interpretable models instead. Nature Machine Intelligence, 1, 206–215.
11. Wahba, G. (1990). Spline models for observational data. SIAM.
12. Wilson, A. G., Hu, Z., Salakhutdinov, R., & Xing, E. P. (2016). Deep kernel learning. In Proceedings of the 19th International Conference on Artificial Intelligence and Statistics (AISTATS) (pp. 370–378).
13. Xu, J., Zhang, X., Liu, J., & Li, H. (2024). AI for social science: A survey of recent developments. Journal of Artificial Intelligence Research, 70, 1–30.
14. Xu, Y., Chen, X., Zhou, Y., & Li, Q. (2019). Deep reinforcement learning for survey design in social science. Proceedings of the AAAI Conference on Artificial Intelligence, 33, 7356–7363.
15. Zhou, D. X. (2018). Deep learning approximation theory. arXiv preprint, arXiv:1805.00631.