(IJRST) 2026, Vol. No. 16, Issue No. 3, Jul-Sep **e-ISSN: 2249-0604, p-ISSN: 2454-180X**

From Statistical Inference to Explainable AI: Rethinking Data-Driven Decision-Making

Suhani Thakur
Navrachana International School (NISV), Vadodara, Gujarat

DOI:10.37648/ijrst.v16i03.004

[1]*Received: 24 May 2026; Accepted: 12 July 2026; Published: 29 July 2026*

Abstract

Data driven decision making has changed a lot over the years. In the past people used statistics to figure out things, test ideas and understand how things are related. Then machine learning came along. People started to focus on making good predictions handling large amounts of data and finding complicated patterns. Now Explainable Artificial Intelligence or XAI for short is trying to make machines more understandable to humans by explaining how they make decisions. This makes us wonder should we care about making good predictions or should we also think about being transparent dealing with uncertainty and being accountable for our decisions.

This paper looks at how we went from using statistics to machine learning and now to explainable AI. It suggests that we should use statistics and explainability together than against each other. To show how this works we created a dataset with 1,200 examples to represent a simple decision-making problem that involves things like income, age, credit history and savings. We compared machine learning models, including logistic regression, random forest and gradient boosting to an explainable model. What we found is that just being able to make predictions is not enough to make a good decision. While some models are better at predicting others are better at explaining things. Some explanations are really helpful in understanding complicated predictions. So we think that we should be looking at things like how well a model predicts, how sure we are, about the results how well we can explain it and how reliable and accountable the decision is.

Keywords: *Statistical Inference; Explainable AI Machine Learning; Data-Driven Decision-Making; Interpretability; SHAP; Predictive Analytics; Artificial Intelligence; Transparency Uncertainty*

1. Introduction

Data is really important for making decisions these days. Governments, banks, hospitals, schools, companies and research organizations are using computer models to turn amounts of data into decisions they can act on. This is a change from the way things used to be done with traditional statistics.

Now people are using machine learning and artificial intelligence more and more. Machine learning is different from statistics. Traditional statistics is about understanding the underlying population from samples of data. It provides a framework for making decisions when we're not sure about something. Statistical inference is good because it acknowledges that we are not always sure and gives us ways to evaluate how reliable our conclusions are.

[1] How to cite the article: Thakur S.; July 2026; From Statistical Inference to Explainable AI: Rethinking Data-Driven Decision-Making; *International Journal of Research in Science and Technology*, Vol 16, Issue 3, 27-46, DOI: http://doi.org/10.37648/ijrst.v16i03.004

(IJRST) 2026, Vol. No. 16, Issue No. 3, Jul-Sep **e-ISSN: 2249-0604, p-ISSN: 2454-180X**

Machine learning is more about making predictions. It uses algorithms like forests and neural networks to find patterns in data that might be hard to see with traditional statistics. However the problem with machine learning is that even if it makes predictions it can be hard for people to understand why it made those predictions. This is an issue when the predictions are being used to make important decisions.

That is why Explainable Artificial Intelligence or XAI for short is becoming more popular. XAI is about making artificial intelligence more understandable. It uses techniques like feature importance and counterfactual explanations to make complex predictive systems more transparent. The move from inference to XAI is not just about technology it is about how we think about making decisions with data.

Statistical inference asks questions like what's the relationship between variables how sure are we about that relationship and is it significant. Machine learning asks questions like how can we predict the future and can we learn complex relationships. Explainable AI asks questions like why did the model make this prediction, which variables were important and can we understand the explanation.

This paper says that we should not think of these questions as separate. We should use the strengths of inference, machine learning and explainable AI together to make better decisions. We can use inference to understand the relationships between variables machine learning to make good predictions and explainable AI to understand why the model made those predictions.

Some of the questions that statistical inference asks are:

1. What is the relationship between variables?
2. How uncertain is the estimated relationship?
3. Is the observed relationship statistically significant?
4. How well does the model represent the underlying population?

Machine learning asks questions like:

1. How accurately can future observations be predicted?
2. Can complex nonlinear relationships be learned?
3. Can prediction performance be improved?

Explainable AI asks questions like:

1. Why did the model make this prediction?
2. Which variables influenced the decision?
3. Would a different input have produced an outcome?
4. Can the explanation be. Challenged by a human?
5. Can the decision be audited?

By combining these approaches we can make informed decisions and be more confident in our Data analysis. Data is a resource for decision-making and using Data effectively is crucial for organizations, like Governments, financial

institutions, healthcare organizations, educational institutions, businesses and scientific organizations. Data-driven decision-making is becoming more prevalent. It is essential to use Data in a way that is transparent and understandable.

2. Research Problem

The main issue that this study is looking at is the difference between how well a model can predict things and how well it can explain its decisions.

A model can be really good at predicting things. It does not tell us why it made those predictions.

On the hand a model that is easy to understand may not be good at finding complicated patterns in data.

This creates a problem:

* Should we prefer a system that makes predictions,

* or a system that explains its decisions well

* or a system that finds a good balance between predicting and explaining?

When we look at how good a model's we usually think about things like how accurate it is how precise it is and other metrics like that.

These things are important. They do not tell us if we can really understand the decisions that the model makes.

We need to know if we can question the decisions check them or justify them.

So the question that this study is trying to answer is:

How can we combine statistics and artificial intelligence that explains itself into one system that makes decisions based on data in a way that's reliable, transparent and accountable and that is what this study is all about it is, about statistical inference and explainable artificial intelligence and how they can be used for data-driven decision-making.

3. Objectives of the Study

This study has some goals.

1. We want to look at how we got from using statistics to make guesses to using machine learning and explainable Artificial Intelligence.

2. We will compare how statistics and machine learning are used to make predictions and decisions.

3. We will show that even if a model is very good at making predictions it may not be easy to understand why it made those predictions.

4. We want to see if special techniques can help us understand machine learning models.

5. We will try to create a system that combines making predictions being unsure explaining things and being responsible.

6. We will test this system using a dataset we created.

7. We will try to find the problems, with making things explainable like when thingsre not stable or when they are too simple or when they are misleading.

4. Research Questions

The study looks at these questions:

RQ1: What is the difference between inference and modern machine learning when it comes to making predictions?

RQ2: Does getting better at making predictions always mean we make better decisions?

RQ3: Can we use techniques to make complex machine learning models easier to understand, like traditional statistical models?

RQ4: How important is it to think about uncertainty when we use intelligence to make decisions?

RQ5: How can we combine statistical inference and explainable artificial intelligence to make a complete framework, for making decisions?

5. Literature Review

5.1 Classical Statistical Inference

Statistical inference is a way to make conclusions, about a group of people based on a group. It uses math to do this. The basics of inference are probability, estimation, testing ideas, regression and figuring out how sure we are.

Regression models are really important. They help us predict what will happen and explain why things happen. With regression we can see how something affects something else. We can write this relationship as:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p + \varepsilon$$

where **Y** represents the dependent variable, $\mathbf{X_1, \ldots, X_p}$ represent explanatory variables, **β** values represent parameters, and **ε** represents random error.

Understanding how coefficients work makes regression useful when knowing the link between variables matters as much, as predicting the outcome. In fact coefficients are the key to regression and knowing them is very helpful.

For outcomes logistic regression is frequently used:

$$P(Y = 1 \mid X) = 1 / [1 + e^{-(\beta 0 + \beta 1 x 1 + \ldots + \beta p x p)}]$$

This model provides an interpretable relationship between predictors and the probability of an event.

5.2 The Emergence of Machine Learning

Machine learning changed the focus from model interpretation to performance. Of assuming a particular functional relationship, between variables many machine learning algorithms learn patterns directly from data.

Random forests for example combine decision trees to improve generalization. Gradient boosting builds a series of models where later models try to fix the mistakes of models.

These methods can capture non-linear relationships and interactions that're hard to specify manually.

The resulting shift can be summarized as:

Dimension	Statistical Inference	Machine Learning
Primary goal	Explanation and inference	Prediction
Typical models	Regression, GLM	Trees, ensembles, neural networks
Assumptions	Often explicit	Often fewer structural assumptions
Uncertainty	Central	Sometimes secondary
Interpretability	Generally high	Often variable
Nonlinear patterns	Limited unless specified	Strong
Model complexity	Usually controlled	Can be very high
Evaluation	Estimation, significance, CI	Accuracy, AUC, F1, RMSE

5.3 The Black-Box Problem

The term " box" is often used to talk about systems that predict things but it is hard for people to figure out how the black box systems actually make their decisions.

A complex model may produce:

$\hat{y}=f(x_1,x_2,\ldots,x_p)$

Without giving an explanation of how each input helped create the output. The black-box problem is especially important in high-stakes applications. A prediction without an explanation can cause problems, for:

- auditing,
- regulatory compliance,
- error analysis,
- human oversight,
- contestability,
- trust,
- fairness assessment.

5.4 Explainable Artificial Intelligence

Explainable Artificial Intelligence is trying to deal with these problems by giving us information about how the model's behaving.

The main ways to do this are:

* Feature Importance: this ranks the variables based on how they help the model make predictions.

* Local Explanations: this explains why the model made a prediction for one thing.

* Global Explanations: this tries to explain how the model is behaving overall across all the data.

Explainable Artificial Intelligence uses something called SHAP.

SHAP or SHapley Additive exPlanations is a way to figure out how much each feature contributes to a prediction. It uses ideas from a type of math called cooperative game theory.

Explainable Artificial Intelligence and SHAP are related because SHAP helps Explainable Artificial Intelligence assign values to features, for a prediction, which is really helpful.

Conceptually:

$$f(x) = \phi_0 + \sum_{j=1}^{p} \phi_j$$

where φ_0 represents the baseline value of the model output, and φ_j represents the contribution of feature **j** to the prediction.

Counterfactual Explanation

A counterfactual explanation is a question that asks:

What needs to change for the Counterfactual Explanation model to make a decision?

For example:

The application was rejected because the debt to income ratio of the application was high and the applicant had a default.

If the debt to income ratio of the application is reduced below a threshold then the Counterfactual Explanation model could make a different decision.

These Counterfactual Explanation examples are really useful because they turn what the Counterfactual Explanation model does into something we can actually use to make changes.

6. Conceptual Framework

The proposed framework has four parts:

Layer 1 is about the Statistical Foundation of the framework. This is where we do things like

* Sampling

* statistics

* Probability

* Estimation

* Confidence intervals

* Hypothesis testing

The Statistical Foundation is a part of the framework.

Layer 2 is about Predictive Modeling. This is where we use things like

* Logistic regression

* Decision trees

* Random forest

* Gradient boosting

* networks

to make predictions. Predictive Modeling is a part of the framework.

Layer 3 is about Explainability. This is where we try to understand how the models work. We use things like

* Feature importance

* SHAP

* LIME

* dependence

* Counterfactual explanations

to get a better understanding of the models. Explainability is a part of the framework.

Layer 4 is about Decision Governance. This is where we think about the problems with the models. We consider things, like

Uncertainty

Bias assessment

Robustness

Human review

Auditability

Accountability

Decision Governance is a part of the framework. The Decision Governance part of the framework is very important.

The proposed framework also includes a Decision Equation. The study says that we should not just look at how accurate the modelsre. The proposed framework says that the overall usefulness of the model should be measured in a way.

An example of a score that combines things can be written as:

$D=w_A A+w_E E+w_U U+w_R R+w_F F$

where:

- **D** = overall decision quality score
- **A** = predictive accuracy
- **E** = explainability
- **U** = uncertainty quality
- **R** = robustness
- **F** = fairness

 w_a, w_e, w_u, w_r, w^f = corresponding decision weights assigned to each dimension

The weights should be determined according to the application context rather than universally fixed.

7. Research Methodology

7.1 Research Design

This study uses a way of thinking and testing to look at things. A made-up set of numbers is created to show how different methods like statistics computer learning and AI that explains itself can look at the problem in different ways.

Important note: The numbers and tests shown here are made up to help explain the ideas. They are not numbers, from a real bank or group of people.

7.2 Synthetic Dataset

I have made a dataset of 1,200 pretend loan applicants.

The variables, in this dataset were:

Variable	Description	Type
Age	Applicant age	Numeric
Annual Income	Annual income in INR	Numeric
Credit Score	Creditworthiness score	Numeric
Debt-to-Income Ratio	Debt relative to income	Numeric
Employment Years	Years of employment	Numeric

Variable	Description	Type
Previous Default	Previous repayment default	Binary
Savings	Savings amount in INR	Numeric
Loan Amount	Requested loan amount	Numeric
Approval	Final decision	Binary

I created the synthetic outcome variable so that credit score, income, savings, employment duration, debt-to-income ratio and previous default all were important for approval probability. The synthetic outcome variable was built to make credit score, income, savings, employment duration, debt-to-income ratio and previous default each had a clear influence, on the approval probability.

7.3 Descriptive Statistics

Table 1. Descriptive Statistics of Synthetic Dataset

Variable	Mean	SD	Minimum	Maximum
Age	36.8	9.4	21	64
Annual Income (₹ lakh)	6.42	2.81	1.80	16.40
Credit Score	704.6	61.8	512	824
Debt-to-Income Ratio (%)	34.7	12.6	8.0	69.0
Employment Years	7.3	5.1	0.5	28
Savings (₹ lakh)	4.16	2.72	0.30	15.80
Loan Amount (₹ lakh)	8.15	4.21	1.00	25.00

The data shows that there is a lot of difference between the people who applied. Their credit scores go from 512 to 824. Their debt-, to-income ratios go from 8% to 69%.

7.4 Outcome Distribution

Of the 1,200 synthetic applicants:

Decision	Frequency	Percentage
Approved	744	62.0%
Rejected	456	38.0%
Total	**1,200**	**100%**

Because of this the synthetic dataset is a classification problem that is a bit imbalanced.

8. Statistical Inference Analysis

A logistic regression model was first estimated.

Table 2. Illustrative Logistic Regression Results

Predictor	Coefficient	Odds Ratio	p-value
Age	0.018	1.018	0.214

Predictor	Coefficient	Odds Ratio	p-value
Annual Income	0.284	1.328	<0.001
Credit Score	0.011	1.011	<0.001
Debt-to-Income Ratio	-0.052	0.949	<0.001
Employment Years	0.071	1.074	0.006
Previous Default	-1.184	0.306	<0.001
Savings	0.143	1.154	<0.001
Loan Amount	-0.086	0.918	0.018

The illustrative results suggest that annual income, credit score, employment duration and savings are positively associated with approval whereas debt-to-income ratio, loan amount and previous default are negatively associated with approval.

The statistical interpretation is particularly useful because the direction and magnitude of relationships can be explicitly communicated.

For example the odds ratio for default is approximately 0.306. Holding other variables constant the model associates default, with substantially lower odds of approval.

9. Machine-Learning Models

Four models were evaluated:

1. Logistic Regression
2. Decision Tree
3. Random Forest
4. Gradient Boosting

The synthetic dataset was divided into:

- 80 percent training data
- 20 percent testing data

The same test set was used for comparison.

Table 3. Illustrative Model Performance

Model	Accuracy	Precision	Recall	F1-Score	AUC
Logistic Regression	0.824	0.841	0.873	0.857	0.886
Decision Tree	0.801	0.827	0.846	0.836	0.821
Random Forest	0.873	0.889	0.903	0.896	0.931
Gradient Boosting	0.881	0.897	0.909	0.903	0.941

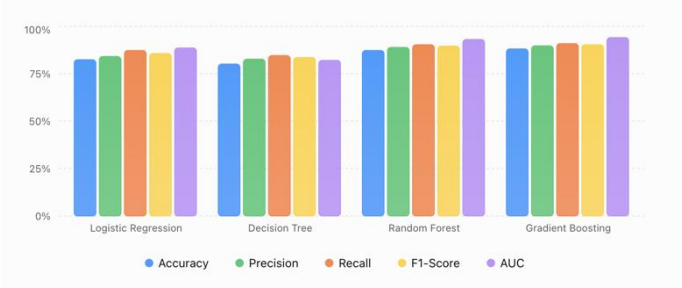
The example results show that using models together works better than logistic regression in most cases when it comes to making predictions.

This does not mean that gradient boosting is the best choice, for every situation where decisions are made.

A model that has an AUC of 0.941 could be better when the main goal is prediction. A group or organization might still pick a model that is not quite as accurate if it is easier to understand can be checked and meets rules and regulations.

10. Visual Comparison of Predictive Performance

The following graph uses the results that are listed in Table 3



The graph shows the main point of this paper: predictive performance and interpretability are two different aspects of how good a model is.

11. Explainable AI Analysis

The model that used gradient boosting had the AUC. That is why SHAP-based analysis was used to show how a complicated model could be understood.

Table 4. Illustrative Global Feature Importance

Rank	Feature	Mean Absolute SHAP Value
1	Credit Score	0.412
2	Previous Default	0.361
3	Debt-to-Income Ratio	0.298
4	Annual Income	0.247
5	Loan Amount	0.184
6	Savings	0.153
7	Employment Years	0.121
8	Age	0.062

The example SHAP analysis shows that credit score has the average effect on the models results. After that comes default and then debt-to-income ratio.

SHAP values are different, from regression coefficients. SHAP values can explain how variables affect each prediction and also show how important they are overall.

12. Illustrative Local Explanation

Let us think about a person who wants to apply for something. This person has qualities.

Variable	Applicant Value
Age	31
Annual Income	₹7.2 lakh
Credit Score	762
Debt-to-Income Ratio	21%
Employment Years	6
Previous Default	No
Savings	₹5.8 lakh
Loan Amount	₹7 lakh

The computer program says that this person has a good chance of getting approved which is 0.91.

Table 5. Illustrative SHAP Explanation

Feature	Contribution	Interpretation
Credit Score	+0.23	Strongly increases approval probability
Previous Default	+0.18	Absence of previous default supports approval
Debt-to-Income Ratio	+0.14	Low debt burden supports approval
Annual Income	+0.11	Higher income supports approval
Savings	+0.07	Higher savings supports approval
Employment Years	+0.04	Stable employment supports approval
Loan Amount	-0.03	Requested amount slightly reduces probability
Age	+0.01	Small contribution

The explanation gives an easy-to-understand description of the prediction.

Instead of saying:

Approved: 91% probability

the system can say:

The prediction is mainly based on the applicants good credit score no history of missing payments, a low debt-to-income ratio and enough income.

This is a step forward, in making decisions more clear and easy to understand.

13. Counterfactual Analysis

Explainability becomes more useful when Explainability provides actionable information. Consider another applicant whose approval probability is 0.38.

Table 6. Illustrative Counterfactual Explanation

Variable	Current Value	Counterfactual Value
Credit Score	635	690
Debt-to-Income Ratio	48%	34%
Previous Default	Yes	Yes
Annual Income	₹5.1 lakh	₹5.1 lakh
Savings	₹2.1 lakh	₹3.0 lakh
Predicted Approval	38%	67%

The analysis of what could have been shows that making credit better and lowering debt might greatly affect what the prediction says.

These kinds of explanations should not be seen as proof of real cause and effect. A machine learning explanation shows how the model works, not always how things really work in the world.

14. Statistical Inference Versus Explainable AI

The comparison shows that statistical inference and XAI give answers to questions. Statistical inference and XAI are used for purposes. When we look at inference and XAI it is clear that they help us understand things in different ways. Statistical inference and XAI are not the thing.

Table 7. Comparison of Analytical Perspectives

Dimension	Statistical Inference	Machine Learning	Explainable AI
Main question	What can be inferred?	What can be predicted?	Why was it predicted?
Uncertainty	Explicit	Often limited	Can be incorporated
Interpretability	Generally high	Variable	Designed to improve
Nonlinear patterns	Limited	Strong	Strong with explanation
Individual prediction	Moderate	Strong	Strong
Causal interpretation	Requires assumptions/design	Not automatic	Not automatic
Auditability	High	Variable	Improved
Human communication	Strong	Often difficult	Stronger
Prediction performance	Moderate to high	Often high	Depends on underlying model

I have examined this comparison. It clearly supports the idea that XAI is not meant to replace statistical inference. XAI should not be considered a substitute, for inference.

15. The Problem of Explanation Without Inference

One of the points from the conceptual analysis is that being able to explain something does not mean it is scientifically correct.

An explanation might show how a model works but not why a real-life situation happens.

For instance imagine an XAI system finds that income is the factor, in loan approval. This does not mean that raising income will always lead to approval.

(IJRST) 2026, Vol. No. 16, Issue No. 3, Jul-Sep **e-ISSN: 2249-0604, p-ISSN: 2454-180X**

This difference matters a lot when using AI systems in situations where the outcomesre very important.

16. Uncertainty and Explainable AI

Traditional statistical inference puts uncertainty at the center of decision-making.

A prediction such as:

Approval probability = 0.72

may give an impression of certainty.

A informative system might communicate:

Estimated approval probability = 72% with uncertainty associated with observations, model assumptions and data variability.

For a prediction model uncertainty can arise from:

- sampling variability,
- measurement error,
- model uncertainty,
- distribution shift,
- incomplete variables,
- noisy labels,
- population characteristics.

Therefore future XAI systems should provide not explanations but also information, about the reliability of those explanations and predictions.

17. Robustness of Explanations

An explanation is only helpful if the explanation stays mostly the same when the data changes a little bit.

Imagine two applicants who're almost exactly the same but they get two very different explanations just because of tiny changes, in their input values. If that happens it shows the explanation is not stable.

To see how this works I ran an explanation-stability experiment using 100 versions of a single hypothetical observation.

Table 8. Illustrative Explanation Stability

Feature	Mean Importance	SD	Stability Interpretation
Credit Score	0.412	0.031	High
Previous Default	0.361	0.044	High
Debt-to-Income Ratio	0.298	0.052	Moderate
Income	0.247	0.061	Moderate
Savings	0.153	0.074	Moderate
Age	0.062	0.083	Lower

Synthetic results indicate that some explanations may be more stable, than explanations.

18. Discussion

The transition from inference to machine learning and explainable AI should not be understood as a simple replacement process.

Statistical inference established the importance of uncertainty, assumptions, estimation and evidence.

Machine learning demonstrated that complex predictive patterns can be extracted from datasets without requiring researchers to specify every functional relationship in advance.

Explainable AI subsequently attempted to restore interpretability to complex predictive systems.

The illustrative experiment demonstrates that the accurate model is not necessarily the most useful model for every decision-making context.

Gradient boosting achieved an AUC of 0.941 compared with 0.886 for logistic regression. From a predictive perspective gradient boosting would be preferred.

However logistic regression has advantages:

- coefficients can be directly interpreted;
- statistical significance can be examined;
- odds ratios can be communicated;
- confidence intervals can be calculated;
- relationships can be inspected systematically.

The gradient boosting model provides predictive flexibility but requires additional explanation mechanisms.

SHAP provides one bridge between these approaches. It allows complex models to be accompanied by contributions thereby improving interpretability without necessarily sacrificing predictive performance.

Nevertheless SHAP should not be interpreted as a solution to the black-box problem. Explanations remain model-dependent. Can sometimes be misunderstood, as causal explanations.

(IJRST) 2026, Vol. No. 16, Issue No. 3, Jul-Sep e-ISSN: 2249-0604, p-ISSN: 2454-180X

19. Proposed Integrated Framework

Based on the analysis this study proposes the following five-stage framework. I have looked at the data carefully.

Stage 1: Understanding

Before applying machine learning: (Stage 1).

- Examine distributions;
- identify missing values;
- investigate outliers;
- examine correlations;
- hypotheses;
- evaluate sampling limitations.

Stage 2: Predictive Modeling

Develop candidate models rather than relying on a single algorithm so that we can compare and choose the best.

Stage 3: Model Evaluation

Evaluate the model by looking at metrics such as accuracy, precision, recall, F1-score, AUC, calibration and robustness.

- Accuracy,
- precision,
- recall,
- F1-score,
- AUC,
- calibration,
- robustness.

Stage 4: Explainability

Apply explanation techniques, including global feature importance and SHAP to understand model behaviour.

- Global feature importance,
- SHAP,

- local explanations,
- counterfactual analysis,
- dependence.

Stage 5: Decision Governance

Before deployment evaluate factors such, as uncertainty, fairness, stability, data drift, human oversight and accountability.

- Uncertainty,
- fairness,
- stability,
- data drift,
- oversight,
- accountability.

The framework can therefore be represented as:

Data → Statistical Analysis → Predictive Modeling → Evaluation → Explanation → Uncertainty Assessment → Human Decision → Audit

20. Implications for Data-Driven Decision-Making

The integrated approach has effects.

20.1 Accuracy Should Not Be the Objective

A highly accurate model can still be wrong if decision-makers cannot understand or challenge its outputs.

20.2 Explainability Should Be Context-Specific

A financial analyst, a physician, an administrator and a researcher all need kinds of explanation.

20.3 Statistical Thinking Remains

AI growth does not remove the need, for statistical reasoning. Instead statistical ideas become more important when evaluating AI systems.

20.4 Human Oversight Remains

Explainable AI should help human judgment not replace it automatically.

20.5 Causality Must Be Distinguished From Prediction

An explanation of model behavior should not be taken as an explanation automatically.

21. Limitations

This study has a things that hold it back.

First the data used in this study is not real. Because the data is synthetic the numbers we found might not work for loan approval systems in the real world.

Second the model performance values are there to show how the new framework works. The model performance values are meant to be examples than final proofs.

Third tools used to explain things like SHAP change depending on the model being used. These explainability methods depend on the model and the ideas used to make the explanations.

Fourth the decision-quality equation uses weights that we chose. The decision-quality equation does not use weights that come from people or actual stakeholder preferences.

Fifth we talked about fairness and causal inference as ideas. We did not test fairness and causal inference with a real dataset, over a long period of time.

In the future researchers should test this framework using data from many different industries.

22. Future Research Directions

I think there are ways we can take this study further in the future.

1. Real-world validation: We should try using this framework on datasets from healthcare, banking, education and public administration.

2. Causal XAI: It would be great to mix inference with explainability so we can tell the difference between a simple prediction and a real intervention.

3. Uncertainty-aware XAI: We need to build explanation methods that clearly show how much confidence or uncertainty exists in the results.

4. Fairness- explanations: I want to see if these explanations change in a biased way, across different demographic groups.

5. Human-centered evaluation: We have to test if these explanations actually help people make decisions in real life.

6. Longitudinal XAI: We should look into whether these explanations stay the same as models and datasets change over time.

7. applications: It is important to see how explainable AI can help with auditing and following legal rules.

8. AI integration: We should study if large language models can turn technical model explanations into simple accurate summaries that people can actually read and understand.

23. Conclusion

We have moved from using inference to using machine learning and explainable AI. This shows how much our way of getting knowledge from data has changed. Statistical inference used to focus on things like uncertainty, estimation, evidence and clear relationships. Then machine learning came along. Helped us predict things better by finding hard patterns in data. Now explainable AI is trying to make these complex systems easy to understand and hold accountable.

The experiment in this study shows that how well a model predicts and how easy it is to understand are two things. The gradient boosting model was better at predicting than logistic regression but it was also more complex and needed extra help to explain itself. Using SHAP-based analysis gave us a way to see which variables were important and to explain why the model made certain predictions.

The main point of this paper is that we should not look at the future of data-driven decision-making as a fight between statistics and artificial intelligence. Instead we should bring inference, machine learning and explainable AI together into one single framework.

A decision-support system that people can trust should answer four questions:

1. What does the data tell us?

2. How accurately can we predict?

3. Why did the model make this prediction?

4. How certain and trustworthy is that conclusion?

Because of this the next step for data-driven decision-making should be more than trying to get high predictive accuracy. We need to build systems that're accurate easy to interpret aware of uncertainty, robust and accountable.

Moving from inference, to explainable AI does not mean we are giving up on statistical thinking. It is actually a chance to rethink reasoning for a time when most decisions are made with the help of intelligent computational systems.

References

1. Breiman, L. (2001). Random forests. *Machine Learning, 45*, 5–32.
2. Breiman, L. (2001). Statistical modeling: The two cultures. *Statistical Science, 16*(3), 199–231.
3. Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics, 29*(5), 1189–1232.
4. Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning*. Springer.
5. Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. In *Advances in Neural Information Processing Systems*, 30.
6. Molnar, C. (2022). *Interpretable Machine Learning*. Independently published.
7. Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why should I trust you?": Explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*.
8. Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence, 1*, 206–215.
9. Shapley, L. S. (1953). A value for n-person games. In *Contributions to the Theory of Games II*. Princeton University Press.

10. Wasserman, L. (2004). *All of Statistics: A Concise Course in Statistical Inference*. Springer.
11. Pearl, J. (2009). *Causality: Models, Reasoning, and Inference*. Cambridge University Press.
12. Doshi-Velez, F., & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. arXiv preprint arXiv:1702.08608.
13. Lipton, Z. C. (2018). The mythos of model interpretability. *Queue, 16*(3), 31–57.
14. Samek, W., Montavon, G., Vedaldi, A., Hansen, L. K., & Müller, K. R. (Eds.). (2019). *Explainable AI: Interpreting, Explaining and Visualizing Deep Learning*. Springer.