IJIRSET
International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY

e-ISSN: 2319-8753 | p-ISSN: 2347-6710


# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

## IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 12, Issue 6, June 2023

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.423

9940 572 462 | 6381 907 438 | ijirset@gmail.com | www.ijirset.com

International Journal of Innovative Research in Science, Engineering and Technology (IJIRSET)

|e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.423| A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 6, June 2023 ||

| DOI:10.15680/IJIRSET.2023.1206155 |

# Evaluating Transparency and Explainability in AI-Driven Planning and Scheduling: A Comprehensive Literature Review

**Prof. Neha Thakre[1], Prof. Ranu Sahu[2], Prof. Kuldeep Soni[3], Deepshika Bisen[4], Anshika Sahu[5]**

Department of Computer Science & Engineering, Badreia Global Institute of Engineering & Management, Jabalpur, M.P., India[1,2,3,4,5]

**ABSTRACT:** As AI technologies rapidly integrate into our work environments, the need for systems that effectively collaborate with humans grows. A crucial aspect of this interaction is the explainability of AI behavior to human users. While AI has surpassed human decision-making abilities in many areas, its opaque algorithms often produce accurate yet incomprehensible results. This lack of transparency can hinder trust between humans and machines.

The TUPLES project, a three-year Horizon Europe research initiative, seeks to address this by developing AI-based planning and scheduling tools with a human-centered focus. By utilizing both data-driven and symbolic AI methods, TUPLES aims to create scalable, transparent, robust, and secure systems. The project employs a use-case-oriented approach, drawing from industry expertise to ensure relevance in areas such as manufacturing, aviation, and waste management.

The EU guidelines for Trustworthy AI emphasize key principles like human oversight, transparency, fairness, societal well-being, and accountability. The Assessment List for Trustworthy Artificial Intelligence (ALTAI) serves as a self-assessment tool for evaluating AI systems. Current planning and scheduling tools only partially meet these criteria, highlighting the need for innovative AI development.

Our literature review explored the transparency and explainability of AI algorithms in planning and scheduling, identifying metrics and recommendations. The study underscores the role of Explainable AI (XAI) in addressing the black box problem, using techniques that make AI systems more understandable and accurate. XAI relies on human explanation concepts to ensure a human-centered approach, offering specific methods and guidance for selecting appropriate tools in planning and scheduling contexts.

**KEYWORDS**: Human-AI interaction, Trustworthy AI, Decision-making

## I. INTRODUCTION

As AI technologies rapidly integrate into professional environments, the development of seamless human-AI collaboration systems becomes increasingly essential. A critical factor for effective human-AI partnerships is ensuring that AI systems exhibit explainable behaviors to humans involved in AI-based decision-making processes. Explainability, also known as interpretable or understandable AI, enables humans to understand the reasoning behind a model's decisions, predictions, or outputs. According to Dwivedi et al. (2023), explainability encompasses concepts such as transparency, which is the AI system's ability to provide clear and comprehensible justifications for its actions; trustworthiness, which refers to the reliability and accuracy of an AI system's decisions; and interpretability, which involves explaining the reasoning behind specific decisions or outputs.

In many domains, AI decision-making has surpassed human capabilities (Minh et al. 2022). However, the opacity of black-box algorithms and the lack of system transparency often result in incomprehensible outcomes (Chromik and Butz, 2021). These opaque results create challenges such as distrust, potential biases, and hindered AI adoption in critical decision-making contexts. Ensuring transparency and explainability in AI systems is crucial to address these issues, fostering trust between humans and machines and enabling users to identify and mitigate biases in AI-generated decisions (Chatila et al. 2021).

Research on Explainable AI (XAI) has made significant progress in addressing the need to make opaque models more transparent (Dwivedi et al. 2023). While many studies have shown that XAI methodologies can improve users'

|e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.423| A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 6, June 2023 ||

| DOI:10.15680/IJIRSET.2023.1206155 |

understanding of black-box models (Adadi and Berrada, 2018; Miller, 2019), others have pointed out the challenges due to the mismatch between individuals' cognitive limitations and current XAI techniques (Bertrand et al. 2022; Bunt, Lount, and Lauzon, 2012). Consequently, there is an increasing need to clarify AI decision-making processes, as a transparent procedure can strengthen trust between humans and machines.

## II. THE TUPLES PROJECT

The European Commission (EC) policy on Trustworthy AI emphasizes creating and testing AI systems that are ethical, transparent, and accountable (European Commission, 2021). These systems must respect human rights, democratic values, and the rule of law, prioritizing human agency and oversight. The policy advocates for AI technologies that enhance societal well-being, ensure privacy, and foster security. Additionally, the EC stresses the need for robustness, fairness, non-discrimination, traceability, and clear communication of AI-driven decisions. These principles aim to build a human-centric AI ecosystem that maintains public trust and promotes sustainable innovation.

The TUPLES project, a three-year research and innovation initiative funded by Horizon Europe, focuses on developing AI-driven planning and scheduling (P&S) tools that emphasize explainability in decision-making processes (TUPLES, 2023). These AI-driven P&S systems aim to optimize resource allocation, streamline processes, and enhance decision-making in complex, dynamic environments. They utilize AI algorithms and techniques to efficiently plan tasks, allocate resources, and schedule activities, considering various constraints, objectives, and uncertainties. By automating the P&S process, these systems can significantly boost productivity, reduce operational costs, minimize human error, and help organizations adapt to changing circumstances (Amershi et al. 2019). Such systems are applicable in a wide range of industries, including manufacturing, logistics, transportation, healthcare, and project management (Kotriwala et al. 2021).

At the heart of the TUPLES project is a comprehensive, human-centric methodology that employs a use-case-oriented approach to ensure practical applicability in real-world scenarios. The project selects use cases based on input from industry experts, cutting-edge advances, and manageable risks, with industries such as manufacturing, aviation, and waste management being key examples. In these various industries, explainable AI in P&S systems can improve transparency and enhance collaboration. For example, AI-driven workforce scheduling systems, which assign employees to shifts based on factors like employee preferences, skill sets, and regulatory requirements, can benefit from explainable AI (XAI). By clarifying the decision-making process, XAI enables managers to adjust schedules based on their understanding of the system's rationale, promoting more effective collaboration.

## III. METHOD

A narrative literature review was carried out to explore techniques and methods for designing and implementing Explainable AI (XAI) systems in planning and scheduling (P&S). The review aimed to identify relevant metrics for assessing XAI and offer recommendations to enhance the explainability and transparency of AI algorithms used in P&S contexts. The literature review involved an extensive search using keywords such as Explainable AI, XAI, Transparency, Trustworthiness, Planning, and Scheduling across bibliographical databases including Scopus, Web of Science, and Google Scholar. This search produced 57 records, which were then evaluated for relevance by the research team. Initially, two authors reviewed the sources based on titles, abstracts, and keywords. Subsequently, two other researchers performed a more detailed examination of the papers' contents to determine their relevance to the review's objectives, outcomes, and recentness. This process was followed by a final round of refinement by two additional researchers based on further criteria. Ultimately, 28 sources were selected for inclusion in the review.

## IV. RESULTS

### IV-A. Techniques and Methods

To advance transparency in AI systems, Minh et al. (2022) conducted a thorough review, categorizing Explainable AI (XAI) methods into three main types: pre-modeling explainability, interpretable models, and post-modeling explainability. Each method has its own set of benefits and drawbacks.

**IV-B. Pre-modeling explainability** focuses on enhancing transparency before the model is trained. This approach aims to identify potential biases and errors early, but it may require extra resources and could complicate the model.

|e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.423| A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 6, June 2023 ||

| DOI:10.15680/IJIRSET.2023.1206155 |

**Interpretable models** are designed to be inherently transparent, offering benefits such as ease of understanding and trust. However, these models might not perform as well as more complex ones and may not be suitable for all applications (Mittelstadt, Russell, and Wachteret, 2019).

**IV-C. Post-modeling explainability** methods work to improve transparency after the model is trained, either through data-driven techniques like artificial neural networks (ANN) or more human-guided methods. Although these methods provide flexibility and can be used with existing models, they may not fully explain the model's decision-making process (Minh et al. 2022).

Several techniques exist to help humans understand complex AI models and their decision-making processes, thereby enhancing explainability (Dwivedi et al. 2023). For instance:

1. **Decision Trees** use a tree-like structure to represent decisions and their possible outcomes, allowing for intuitive visualization and interpretation (Mahbooba et al. 2021).
2. **Fuzzy Logic** involves reasoning with approximate values rather than strict true or false values, facilitating more human-like reasoning and interpretation of AI logic (Chimatapu et al. 2018).
3. **Bayesian Networks** are graphical models that illustrate probabilistic relationships among variables, offering insights into causal dependencies and probabilistic reasoning (Hennessy, Diz, and Reiter, 2020).

**Additionally:**

1. **Local Interpretable Model-Agnostic Explanations (LIME)** generates post-hoc explanations for individual predictions by approximating complex model behavior with a simpler, interpretable model in the vicinity of a specific prediction (Upadhyay et al. 2021).
2. **Counterfactuals** explore alternative scenarios to understand AI decision-making by answering "what-if" questions and showing how changes in input features could lead to different outcomes (Byrne, 2019).
3. **SHAP (SHapley Additive exPlanations)** calculates Shapley values based on cooperative game theory to assign importance to features in machine learning models. This approach enhances transparency and interpretability by revealing the impact of individual features on model predictions (Chromik, 2020).

## V. METRICS FOR XAI EVALUATION

The literature identifies several crucial metrics for evaluating the explainability of AI systems, which can be adapted for planning and scheduling (P&S) contexts. Sovrano et al. (2022) categorize these metrics into three main classes:

1. Model-Based Metrics: These include accuracy, precision, and recall.
2. Post-Hoc Metrics: Techniques such as LIME and SHAP fall into this category.
3. Subject-Based Metrics: These involve human subjects to assess explainability through methods like surveys and interviews.

According to Sovrano et al. (2022), various metrics can complement each other and serve distinct purposes based on the context.

|e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.423| A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 6, June 2023 ||

| DOI:10.15680/IJIRSET.2023.1206155 |

**Table 1.** Potential of XAI techniques in TUPLES use cases.

| XAI Technique | TUPLES use case 1: Pilot assistance | TUPLES use case 2: Aircraft manufacturing |
|---|---|---|
| Decision Tree | It helps the flight pilot decide where to land the plane in case of an emergency, based on factors such as weather conditions and airport location. | It evaluates worker skills, material needs, and task order, presenting a step-by-step guide to optimally assign resources and workers to tasks. |
| Fuzzy Logic | It helps the pilot decide when there are ambiguous or conflicting data inputs, such as the severity of the emergency or the status of the plane's systems. | It assists the manager in deciding where to allocate multi-skilled workers for tasks, while addressing uncertain worker skill levels and fluctuating resource constraints. |
| Bayesian Networks | It helps the pilot decide by weighing the probabilities of different outcomes, such as the likelihood of a successful landing at a particular airport. | It facilitates managers' decisions by estimating success probabilities of tasks, incorporating uncertainties such as worker experience or resource availability. |
| LIME | It helps the pilot explain its decision-making process to human pilots and passengers, by highlighting which data inputs were most influential in the decision. | It clarifies AI-recommended worker allocations by highlighting influential factors like experience, task urgency, and resource constraints, in aircraft assembly scheduling. |
| Counterfactual | It helps the pilot learn from previous decisions by exploring what might have happened if it had chosen a different landing site or approach. | It demonstrates alternative workforce allocations to managers by comparing expected outcomes and delays to the AI-recommended decisions. |
| SHAP | It assists pilots in understanding how various factors, such as aircraft speed and altitude, contribute to aircraft fuel consumption, by assigning each factor a numerical importance score. | It provides managers with numerical scores that prioritize factors like worker skills and resource availability for workforce allocation decisions in tasks such as avionics installation or wing assembly. |

## VI. KEY FACTORS AND APPROACHES FOR EVALUATING AI SYSTEM EXPLAINABILITY INCLUDE

1. Simplicity: This measures the complexity or straightforwardness of the explanations provided. Simpler explanations are generally easier for users to understand (Miller, 2019).
2. Fidelity: This metric assesses the accuracy of the explanations in reflecting the underlying AI model's behavior. High fidelity means the explanations accurately describe the model's decision-making process (Velmurugan et al. 2021).
3. Consistency: This refers to the stability of explanations when small changes are made to the input or the model. Consistent AI systems offer similar explanations despite minor alterations (Hoffman, 2018).
4. Transparency: This metric evaluates how much insight the AI system provides into its internal workings. More transparent systems offer greater visibility into their decision-making processes (Felzmann et al. 2020).
5. Local Explanations: These explain the decision-making process for specific instances.
6. Global Explanations: These provide a broader understanding of the AI system's behavior across multiple instances (Setzu et al. 2021).
7. Actionability: This measures the usefulness of the explanations for users to make informed decisions or take actions based on the AI system's output. Actionable explanations guide users on modifying input features to achieve desired outcomes (Joshi et al. 2019).

Evaluating AI systems' explainability in P&S contexts requires considering a range of metrics. Each metric serves a distinct purpose, and their combination provides a comprehensive assessment of the system's ability to convey its decision-making process effectively (Sovrano et al. 2022). Selecting the appropriate metrics based on specific use cases

|e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.423| A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 6, June 2023 ||

| DOI:10.15680/IJIRSET.2023.1206155 |

and user needs is essential for ensuring meaningful and practical insights, ultimately fostering trust and collaboration between humans and AI systems.

## VII.BENEFITS OF EXPLAINABILITY

Users who grasp how an AI system makes decisions are more likely to trust and accept its results (Chatila et al. 2021). This trust is essential for widespread technology adoption across different industries. Adadi and Berrada (2018) highlighted that Explainable AI (XAI) has considerable potential to improve trust and transparency in AI systems, which is especially crucial in sectors such as healthcare, finance, and law enforcement, where AI-driven decisions can significantly impact individuals and society. Kotriwala et al. (2021) also emphasized that transparent explanations from XAI tools are vital for fostering trust between humans and AI in these critical areas. However, implementing XAI in the process industry presents challenges due to the varied needs of different AI end-users and application scenarios (Kotriwala et al. 2021)

When stakeholders can comprehend the reasoning behind AI-generated insights, they are better equipped to make informed decisions, leading to improved outcomes. Cartovloni et al. (2022) examined the ethical, legal, and social aspects of AI-based medical decision-support tools and found that XAI could offer clear explanations of AI system actions and decisions. This transparency helps patients understand the decision-making process and evaluate if it meets ethical and regulatory standards. By elucidating algorithmic decisions, XAI systems assist users in understanding the factors influencing outcomes and support effective debugging, which is particularly valuable in complex fields like planning and scheduling (P&S), where errors can have significant impacts (Schmid and Wrede, 2022).

## VIII. TRADE-OFF BETWEEN ACCURACY AND EXPLAINABILITY

The trade-off between accuracy and explainability in AI poses a significant challenge, as it involves finding a balance between a model's predictive performance and its interpretability. Gunning and Aha (2019) highlight an inherent conflict between these two aspects: the most accurate models are often the least explainable, while the most explainable models usually offer lower accuracy.

This trade-off manifests in two distinct types of models:

1. Complex Models: For example, deep neural networks are known for their high accuracy but low explainability. They achieve impressive predictive performance but are difficult for users to interpret, making it challenging to understand how they arrive at their predictions.
2. Simpler Models: Such as decision trees, which offer high explainability but lower accuracy. These models are easier for humans to understand and can be visually represented, providing a clear decision-making process. However, their simplicity means they may not capture complex data relationships as effectively, leading to lower predictive accuracy compared to more complex models (Gunning and Aha, 2019).

Therefore, finding an optimal balance between accuracy and explainability is essential for developing AI systems that not only enhance human decision-making but also ensure trust and transparency in their operations.

### VIII: A. Human-Centered Approaches

Human-centered approaches to Explainable AI (XAI) can lead to more adaptable and flexible systems that meet users' varied needs and preferences. Chromik and Butz (2021) argue that explanation interfaces should allow users to adjust the level of detail and complexity of explanations according to their needs. Different users have varying requirements for explanations based on their skills, expertise, and context.

Tailoring AI explanations to individual users can make explanatory tools more useful across a diverse range of users. Miller (2019) highlights the importance of understanding human explanations to improve AI explanations. He suggests that AI can benefit from existing models of how people formulate and present explanations. Human explanations often address "why" questions contrastively, are selectively biased, have a social dimension, and prioritize causal links over probabilities. Integrating these aspects into AI models can enhance their explanatory capabilities, although this may not be applicable in all scenarios.

Bertrand et al. (2022) reviewed 37 papers on cognitive biases in XAI systems and found that these biases can impact the transparency and accuracy of AI decision-making. They identified several techniques to mitigate these biases, such as employing multiple explanation methods and involving users in the design process. Their study produced a heuristic map to guide the development of XAI systems that align better with human cognitive processes. This underscores the

|e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.423| A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 6, June 2023 ||

| DOI:10.15680/IJIRSET.2023.1206155 |

importance of incorporating human factors into the design of XAI systems to improve their effectiveness and acceptance.

### VIII-B. Interdisciplinary Approach

Interdisciplinary collaboration is essential for advancing research in Explainable AI (XAI), integrating expertise from fields such as data science, human-computer interaction, cognitive psychology, and philosophy (Adadi and Berrada, 2018). Incorporating human factors into the design and development of XAI systems is vital, as cognitive biases can impact their transparency and accuracy (Bertrand et al., 2022). Schmid and Wrede (2022) emphasize the importance of interdisciplinary and linguistic expertise in creating dialog modeling approaches for adaptive, multi-modal interactions. They stress that explanations need to be tailored to individual user needs and align with human cognitive processes.
Chromik and Butz (2021) examine the interdisciplinary nature of XAI, which seeks to improve human understanding of complex machine-learning models through explanation methods. They highlight "flexibility" as a key design principle for interactive explanation interfaces in XAI. Such interfaces should allow users to adjust the detail and complexity of explanations to meet their specific needs and preferences, thereby supporting informed decision-making by enhancing their comprehension of the AI system's behavior.

## X. CONCLUSION

This paper offers a thorough review of Explainable AI (XAI) systems within the Planning and Scheduling (P&S) domain, highlighting the essential roles of transparency, explainability, and interdisciplinary collaboration in advancing AI research. It identifies various XAI methods and provides evidence to guide the selection of appropriate XAI tools for P&S, revealing numerous opportunities to enhance system effectiveness and user acceptance.

One significant opportunity is to bolster trust and transparency, which are critical for the continued development of AI systems (Adadi and Berrada, 2018). In fields like healthcare, where AI-driven decisions can greatly affect patient treatments, XAI helps users understand the AI's decision-making process, thereby increasing trust and transparency. This, in turn, leads to more informed decision-making and greater confidence in AI recommendations.

## ACKNOWLEDGEMENT

The findings suggest several recommendations for effectively implementing XAI in P&S. First, human-centered design should be prioritized by involving end-users in the development process to ensure that explanations are actionable and informative (Chromik and Butz, 2021). Second, fostering interdisciplinary collaboration among experts from data science, cognitive psychology, and philosophy is crucial (Schmid & Wrede, 2022). Third, there should be a focus on identifying and standardizing evaluation metrics for XAI in P&S systems to objectively measure their impact. Finally, addressing cognitive biases in XAI systems is essential to improve transparency and accuracy (Chromik and Butz, 2021).

Future research should investigate adaptive XAI systems and evaluate their effect on user trust and adoption across various P&S scenarios. Aligning AI systems with human cognitive processes and customizing explanations to users will enhance trust, improve decision-making, and address ethical concerns, ultimately supporting the broader adoption of AI in the P&S domain.

## REFERENCES

1. Adadi, A., Berrada, M. (2018) "Peeking inside the black-box: a survey on explainable artificial intelligence (XAI)". IEEE access, Volume 6, pp. 52138–52160.
2. Amershi, S., Weld, D., Vorvoreanu, M., Fourney, A., Nushi, B., Collisson, Suh, J., Iqbal, et al. (2019) "Guidelines for human-AI interaction", Proceedings of the CHI Conference on Human Factors in Computing Systems, Glasgow, Scotland, UK.
3. Bertrand, A., Belloum, R., Eagan, J. R. and Maxwell, W. (2022) "How Cognitive Biases Affect XAI-assisted Decision-making: A Systematic Review", Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society (AIES'22), Oxford, UK.
4. Bunt, A., Lount, M., and Lauzon, C. (2012) "Are explanations always important? A study of deployed, low-cost intelligent interactive systems", Proceedings of the 2012 ACM international conference on Intelligent User Interfaces, Lisbon, Portugal.

|e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.423| A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 6, June 2023 ||

| DOI:10.15680/IJIRSET.2023.1206155 |

5. Byrne, R. M. (2019) "Counterfactuals in Explainable Artificial Intelligence (XAI): Evidence from Human Reasoning", Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence (IJCAI-19), Macao, PRC.
6. Cartolovni, A., Tomi ˇ ci ˇ c, A., and Mosler, E. L. (2022) "Ethical, legal, and social ´ considerations of AI-based medical decision-support tools: A scoping review", International Journal of Medical Informatics, Volume 161, no. 104738.
7. Chatila, R., Dignum, V., Fisher, M., Giannotti, F., Morik, K., Russell, S., and Yeung, K., eds. (2021). Trustworthy AI, Reflections on Artificial Intelligence for Humanity, Lecture Notes in Computer Science. Springer, Cham. pp. 13–39.
8. Chimatapu, R., Hagras, H., Starkey, A., and Owusu, G. (2018), "Explainable AI and fuzzy logic systems", Proceeding of Theory and Practice of Natural Computing: 7th International Conference, Dublin, Ireland, Springer International Publishing, pp. 3–20.
9. Chromik, M. (2020) "Reshape: A framework for interactive explanations in xai based on shap", Proceedings of 18th European Conference on ComputerSupported Cooperative Work, Siegen, Germany, European Society for Socially Embedded Technologies, Volume 4, no. 2, pp 1–5.
10. Chromik, M., Butz, A. (2021) "Human-xai interaction: A review and design principles for explanation user interfaces", in Proceedings of 13◦ International Conference Human-Computer Interaction, Bari, Italy, Springer International Publishing, pp. 279–298.
11. Dwivedi, R., Dave, D., Naik, H., Singhal, S., Omer, R., Patel, P., Quian, B., Wen, Z., et al. (2023) "Explainable AI (XAI): Core ideas, techniques, and solutions", ACM Computing Surveys, Volume 55, no. 9, pp. 1–33.
12. European Commission, High-level expert group on artificial intelligence (2021). Deliverable 1: Ethics guidelines for trustworthy AI. https://digital-strategy.ec.eu ropa.eu/en/policies/expert-group-ai.
13. Felzmann, H., Fosch-Villaronga, E., Lutz, C., and Tamò-Larrieux, A. (2020) "Towards transparency by design for artificial intelligence", Science and Engineering Ethics, Volume 26, no. 6, pp. 3333–3361.
14. Gunning, D., and Aha, D. W. (2019) "DARPA's explainable artificial intelligence (XAI) program", AI Magazine, Volume 40, no. 2, pp. 44–58.
15. Hennessy, C., Diz, A. B. and Reiter, E. (2020) "Explaining Bayesian Networks in natural language: state of the art and challenges" 2nd Workshop on Interactive Natural Language Technology for Explainable Artificial Intelligence, pp. 28–33.
16. Hoffman, R. R., Miller, T., Mueller, S. T., Klein, G., and Clancey, W. J. (2018) "Explaining explanation, Part 4: A deep dive on deep nets", IEEE Intelligent Systems, Volume 33, no. 3, pp. 87–95.
17. Joshi, S., Koyejo, O., Vijitbenjaronk, W., Kim B. and Ghosh, J. (2019) "Towards realistic individual recourse and actionable explanations in black-box decision making systems" arXiv:1907.09615: https://arxiv.org/abs/1907.09615
18. Kotriwala, A., Klöpper, B., Dix, M., Gopalakrishnan, G., Ziobro, D. and Potschka, A. (2021) "XA`I for Operations in the Process Industry-Applications, Theses, and Research Directions", Proceedings of the AAAI 2021 Spring Symposium on Combining Machine Learning and Knowledge Engineering, California, USA, Volume 2845, no. 26.

INNO SPACE
SJIF Scientific Journal Impact Factor

Impact Factor: 8.423

doi cross ref

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर NISCAIR

# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

## IN SCIENCE | ENGINEERING | TECHNOLOGY

9940 572 462 6381 907 438 ijirset@gmail.com

www.ijirset.com

Scan to save the contact details