International Journal of

Technology and Systems

(IJTS)

Sprint Planning and Al/Ml: How to Balance Iterations with Data Complexity

Journal of Technology and Systems
ISSN : 2788-6344 (Online)
Vol. 6, Issue No. 2, pp. 56 – 72, 2024
www.carijournals.org

Sprint Planning and Al/Ml: How to Balance Iterations with Data Complexity

[1*] Ankur Tak, [2] Sunil Chahal
https://orcid.org/0009-0005-8548-1636

Accepted: 20th Mar, 2024 Received in Revised Form: 5th Apr, 2024 Published: 20th Apr, 2024

Abstract

Purpose: This study aims to bridge the gap in modern software development by integrating agile methodologies with artificial intelligence and machine learning (AI/ML). It seeks to understand how agile sprint planning can effectively interact with the complexities of data inherent in AI/ML projects.

Methodology: The approach taken in this research draws upon a hermeneutic philosophy, which is inference-based, alongside a descriptive methodology. It investigates effective strategies for data preparation, model building, and validation, utilizing the iterative architecture of agile sprinting. The study also incorporates a critical examination to address significant limitations encountered in the integration process.

Findings: The findings highlight the essential role of cross-departmental teams and identify various technological tools that facilitate the smooth integration of agile methodologies with AI/ML projects. A rigorous examination also emphasizes the necessity for ongoing validation through evidence to manage the complexities effectively.

Unique Contribution to Theory, Practice, and Policy (Recommendations): The study offers a comprehensive framework and practical recommendations for businesses aiming to handle data-driven AI/ML initiatives in agile environments. It provides a strategic management approach that aligns more successfully with the demands of data production and agile processes, thus contributing significantly to both theoretical perspectives and practical applications in software development. These contributions are pivotal for informing policy on the integration of cutting-edge technologies in agile settings.

Keywords: *Agile, Neural Networks, Complexity, Machine Learning, Phenomenon, Integration Strategies*

Journal of Technology and Systems
ISSN : 2788-6344 (Online)
Vol. 6, Issue No. 2, pp. 56 – 72, 2024
www.carijournals.org

Introduction

1.1 Research background

The incorporation of agile approaches represents a crucial turning point in the rapidly changing field of machine learning as well as artificial intelligence (AI/ML). Agile, which emphasizes iterative development as well as adaptability, contrasts with AI/ML projects, which are frequently unpredictable as well as data-intensive [1]. The traditional sprint planning method, which was created for clearly defined tasks, might not perfectly match the complicated requirements of AI/ML methods, which heavily depend on the quantity, accuracy, and complexity of the data. Finding a balance between sprints'-controlled rhythm and the fluid requirements of AI/ML development thus becomes a crucial problem [2]. This study aims to explore the frameworks and tactics used to integrate these seemingly unrelated methodologies. This study intends to give practical insights for both individuals and organizations negotiating the complex intersection of sprint organizing and AI/ML projects by examining successful cases, and philosophical frameworks, and developing best practices [3]. In the end, our research aims to contribute to a method for managing AI/ML projects in agile environments that is more successful and efficient.

1.2 Research aim and objectives

Research Aim:

The purpose of this study is to create a thorough framework for efficiently integrating AI/ML applications into agile sprint execution, with data complexity as the main determining factor.

***Objectives*:**

- To evaluate how data complexity affects how AI/ML projects organize their sprints.
- To identify the main obstacles and stumbling blocks in integrating AI/ML development into agile approaches.
- To evaluate the practices currently being used by enterprises to manage data-intensive AI/ML projects during sprints.
- A set of principles and best practices for optimizing time management in the environment of AI/ML assignments, taking different levels of knowledge complexity into consideration.

1.3: Research Rationale

In modern software development, it is crucial to integrate agile approaches with neural networks and machine learning (AI/ML). Agile's iterative methodology clashes with the complicated requirements of AI/ML projects, which are highly dependent on data complexity as well as quality [4]. Effective sprint management becomes challenging as a result of this incongruity. The key to speeding development procedures in AI/ML applications is to comprehend how data complexity affects sprint planning [5]. By offering useful insights and a systematic methodology, this research aims to close this gap and improve the effectiveness and profitability of AI/ML projects inside agile organizations [6]. For industries that have substantially committed in based on artificial intelligence solutions, these developments have far-reaching ramifications.

2.1: Foundations of Agile Software Development

The tenets and procedures that make up the Agile Software Development foundations are all geared toward encouraging adaptability, teamwork, and response in the computer program creation process [7]. Agile is a methodology that first gained popularity in the first decade of the 2000s as a reaction against the rigidity of conventional waterfall techniques. Its key principles, as stated in the Rapid Manifesto, give importance to people and interactions, practical solutions, customer communication, and change-responsiveness [8]. Iterative processes of development, wherein brief timeframes (sprints) are used to produce tiny amounts of functionality, combined ongoing input loops, which enable quick adaptability to changing needs, are key elements of Agile. Teams from various departments also cooperate while fostering mutual responsibility and erasing traditional role distinctions. Specific frameworks for putting these concepts into practice are provided by agile approaches like Scrum, and Kanban, especially Extreme Programming (XP) [9]. For example, Scrum uses weekly stand-up sessions, and time-limited iterations, including sprint reviews to enable regular progress evaluation.

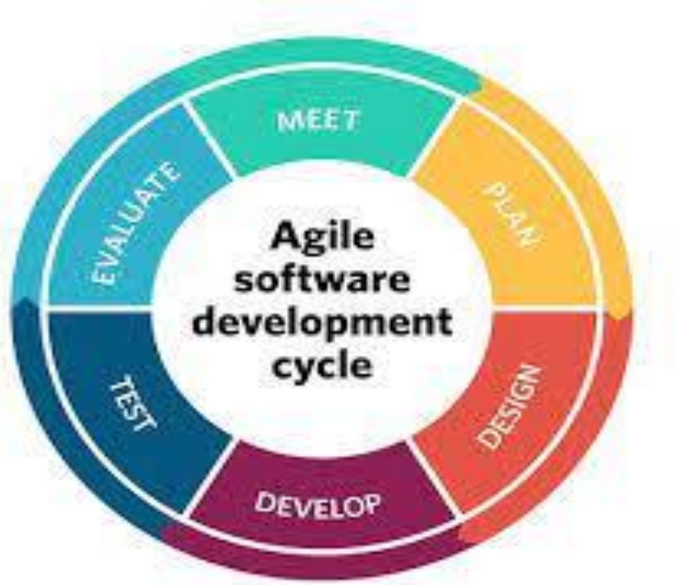

Fig. 1: Agile Development Method

To maximize productivity, Kanban focuses on monitoring workflow including limiting tasks to progress. To assure high-quality results, XP emphasizes techniques like test-driven design and pair programming [10]. Overall, the principles of agile software construction transformed

Journal of Technology and Systems
ISSN : 2788-6344 (Online)
Vol. 6, Issue No. 2, pp. 56 – 72, 2024

project management by emphasizing adaptability, customer service, and iterative progress, which allowed teams to more effectively navigate challenging and changing development environments.

2.2: Complexity in AI/ML Projects: Data Considerations

The complex makeup of data is largely what gives AI/ML projects their level of sophistication. Data concerns are crucial since they determine whether machine instruction and artificial intelligence projects succeed or fail [11]. Data amount, variety, velocity, and authenticity are factors that affect complexity. Volume is the total amount of data, and can range from enormous datasets in large-scale data applications to smaller datasets in specialized fields. Data can take many different forms, including data that is organized, semi-organized, and unorganized, hence a variety of processing methods are required [12]. Velocity is the pace at which information is produced, necessitating immediate processing for some applications. Veracity examines the dependability and correctness of data, an important aspect of developing precise models. The history of information and provenance are also essential for assuring ethical and legal compliance.

Fig. 2: Managing Machine Learning

Strong data pretreatment, transformation, and cleaning processes are required to handle complicated data, frequently requiring methods like feature extraction and reducing dimensionality [13]. Furthermore, comprehension of the fundamental information's complexity is necessary for choosing the right machine-learning methods and models. Therefore, a key component of the accomplishment of AI/ML projects is understanding and successfully managing data complexity [14]. It involves a comprehensive strategy that incorporates data engineering, preliminary processing, including modeling techniques customized to the unique properties of the relevant dataset.

2.3: Integrating AI/ML with Agile Methodologies: A Comparative Analysis

ISSN : 2788-6344 (Online)
Vol. 6, Issue No. 2, pp. 56 – 72, 2024

Given the difficulties presented by data-driven initiatives, integrating AI/ML alongside Agile Methodologies calls for a sophisticated approach. This comparative research examines the overlaps and differences between the sophisticated, data-centric character of AI/ML development and Agile's iterative, adaptive methodology [15]. Although both approaches place a high priority on adaptability, Agile's strict sprint schedules might not be compatible with AI/ML's fluctuating data needs. As data preprocessing including model training are crucial but frequently invisible aspects in AI/ML projects, the agile commitment to software that works may be hard. Case studies of effective integration are useful standards [16]. They emphasize tactics like using teams from different departments with data knowledge and adding feedback loops for bettering data quality.

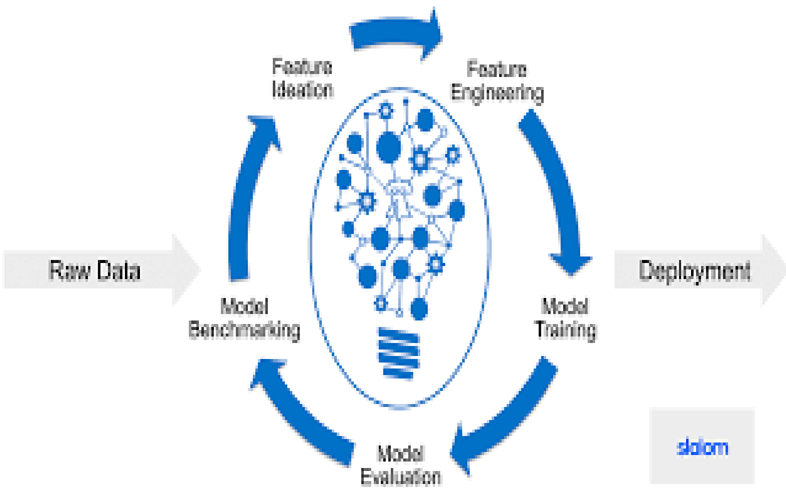

Fig. 3

Another way to close the gap is to implement iterative milestones within sprints expressly for data-related activities [17]. In the end, this analysis tries to glean takeaways for integrating Agile with AI/ML, acknowledging the importance of a customized, adaptable strategy for maximizing the possible benefits of projects based on data under Agile frameworks.

2.4: Guidelines and Methodologies for Effective Sprint Planning in Data-Intensive AI/ML Projects

Due to a seamless combination of these intricate areas, guidelines and procedures for effective sprint preparation in heavy-on-data AI/ML programs are essential. First and foremost, careful data preprocessing needs to be given top priority, with time allotted for jobs like cleaning, design of features, and validation [18]. It is essential to clearly define sprint goals, focusing on attainable benchmarks that are consistent with the larger scope of the project. A complete strategy is fostered by creating multidisciplinary groups with various knowledge, including data scientists, technology engineers, as well as domain specialists [19]. It's also crucial to include buffer time to handle unforeseen data complexity or model adjustment. To ensure the reliable performance of the models, rigorous evaluation and validation techniques should be incorporated into the sprint

Journal of Technology and Systems
ISSN : 2788-6344 (Online)
Vol. 6, Issue No. 2, pp. 56 – 72, 2024
www.carijournals.org

schedule [20]. Aligning your sprint's achievements with organizational goals requires ongoing feedback channels with stakeholders.

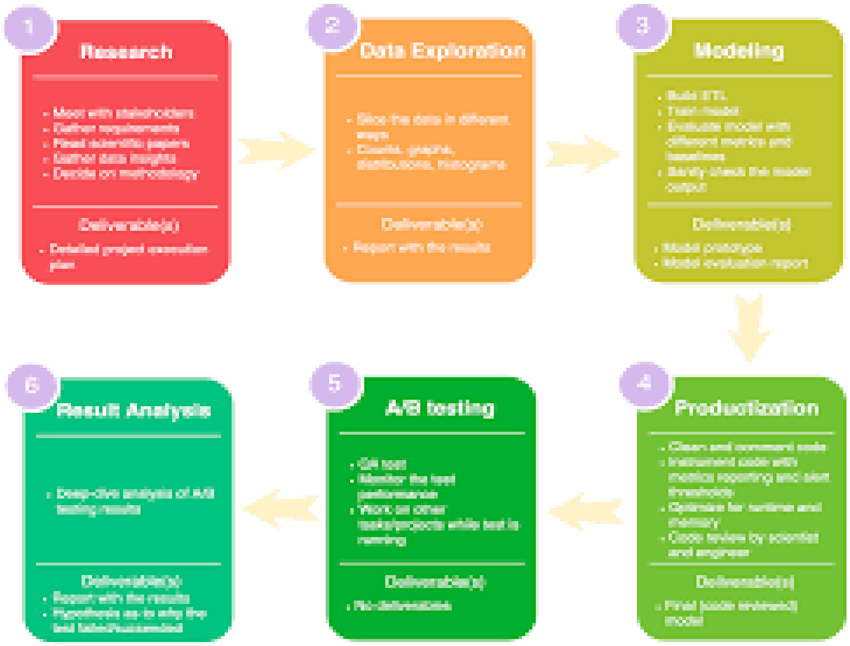

Fig. 4

Agile teams should also use workflow management software and visualization programs to efficiently track data-related operations [21]. Finally, to reflect on experiences learned and improve sprint planning processes for further iterations, regular retrospectives must be held. With this all-encompassing strategy, sprint planning within data-intensive AI/ML applications is not only effective but also conducive to producing high-quality, significant results [22].

2.5: Literature Gap

There is currently a dearth of comprehensive frameworks that handle the unique difficulties provided by complicated data in sprint preparation in the literature on incorporating AI/ML with agile approaches. While studies have looked at certain parts of AI/ML technology with agile approaches, there is a clear need for more information on how to balance iterations with the complexities of heavily data-driven projects under a framework based on agile.

METHODOLOGY

This study is consistent with interpretivism, which acknowledges that AI/ML initiatives under agile organizational structures are influenced by individual interpretations and contextual details. It recognizes that human perception and experiences are crucial in determining how such integrations turn out [23]. The precise insights along with recommendations for merging AI/ML with agile approaches will be derived using a method called deductive reasoning. This entails creating theories-based hypotheses and methodically testing them via empirical observations as

Journal of Technology and Systems
ISSN : 2788-6344 (Online)
Vol. 6, Issue No. 2, pp. 56 – 72, 2024
www.carijournals.org

well as analysis. For this topic, a descriptive research strategy is appropriate [24]. With a focus on heavily data-driven projects, it attempts to give a thorough and comprehensive overview of the state and existing practices of incorporating AI/ML with agile approaches. A thorough understanding of the phenomenon is provided by descriptive research, which aids in capturing the subtleties of how businesses approach this integration. Secondary data sources will be actively utilized in this investigation [25]. To compile pertinent data, a thorough assessment of academic publications, proceedings of conferences, whitepapers, reports from the sector, and case histories will be done. This strategy was used because it effectively gathers a wide range of viewpoints and knowledge from many sources [26]. To identify current frameworks, obstacles, and best practices for incorporating AI/ML with agile approaches, particularly in data-driven projects, a thorough literature study will be done. This will entail looking through Google Scholar, IEEE Xplore, and the ACM Digital Library [27]. The chosen works of literature will be examined for relevance, reliability, and recentness. Only works that have undergone peer review and have recently been published at a conference or through an authoritative assessment will be taken into account. We will extract pertinent information from the chosen literature, such as integration strategies, difficulties encountered, as well as effective implementation examples [28]. Cases that illustrate the handling of information's complexity will receive additional consideration. A theme analysis will be performed on the extracted data. We will look for trends, commonalities, and differences in the methods used to integrate AI/ML with agile [29]. A particular focus will be given to data complexity management techniques. We'll use qualitative analysis of content to examine the data. We'll find commonalities and trends in the integration of AI/ML with agile [30]. Furthermore, particular approaches and methods for dealing with information complexity will be retrieved and categorized. In order to acknowledge the subjective aspect of the procedure of integration, the meanings will be based on the interpretive philosophical framework [31].

RESULTS

4.1 Data Preprocessing Strategies for Agile AI/ML Integration

Effective data preparation is crucial when combining AI/ML with agile approaches. Agile's iterative methodology demands efficient procedures that may easily change to meet changing data needs [32].

***Automated Data Cleaning and Imputation**:*

By using automated instruments and techniques to deal with missing values, unusual values, as well as noise, data is always clean and available for model training [33]. By doing this, the preprocessing stage is sped up, and manual involvement is decreased.

Dynamic Feature Selection and Engineering:

AI/ML projects that are agile can take advantage of feature selection methods that can be adjusted to find and rank pertinent qualities [34]. As the project develops, additional, instructive features can be created using dynamic feature engineering approaches.

Data Normalization and Standardization:

It is essential to standardize data scales as well as distributions so that models may efficiently incorporate knowledge from various aspects [35]. Uniform data representation is made possible by using methods like z-score multiplication or min-max standardization.

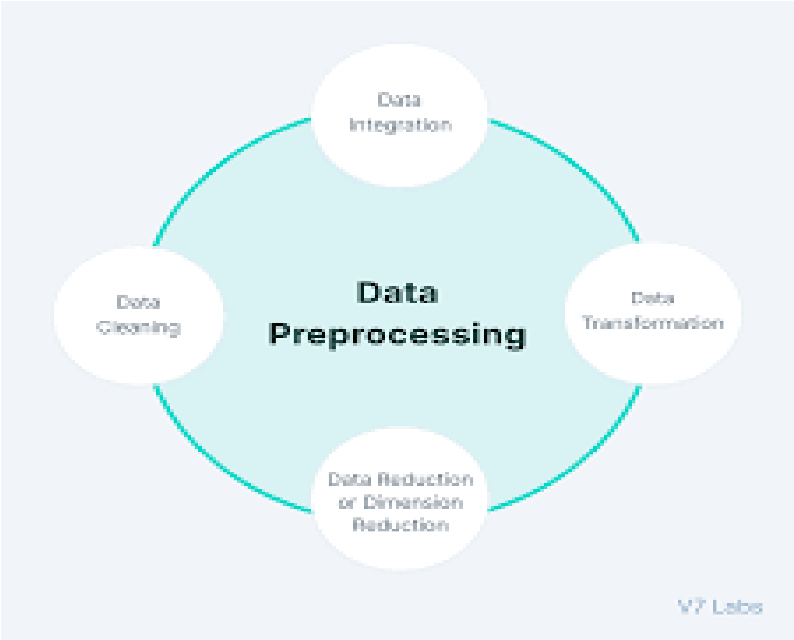

Fig. 5

Incremental Data Updates and Versioning:

In agile contexts, data sets may change throughout the course of sprints. While maintaining the integrity of existing models, version control along with incremental update procedures enable the seamless incorporation of new data [36].

Data Quality Monitoring and Feedback Loops:

Data accuracy and consistency measures, which are continuously monitored, aid in the early detection of problems [37]. It is possible to make quick adjustments based on changing data complexity by establishing a cycle of feedback between data pretreatment and model construction.

4.2 Agile-Compatible Model Development and Validation Techniques

Model creation and validation methods that fit with the incremental and adaptable spirit of nimble sprints are necessary for incorporating AI/ML with agile approaches [38]. These techniques guarantee that models adapt to changing inputs and needs in an efficient manner.

Incremental Model Training:

Models can be trained incrementally so that new data can be incorporated, and current models can be improved rather than starting from scratch with each sprint [39]. The model can adjust to changing patterns and trends thanks to this strategy.

Ensemble Learning and Model Stacking:

Using ensemble techniques like bagging, boosting, and stacking, it is possible to combine various models into a solid, integrated system [40]. This strategy can improve model functionality and resist shifting data dynamics.

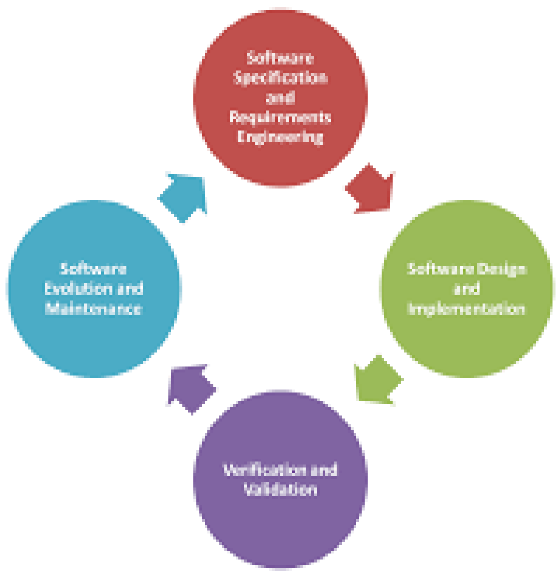

Fig. 6

Cross-Validation Strategies:

Model adaptation across several subsets of data can be evaluated using cross-validation methods that are consistent with agile methodologies, such as k-fold cross-validation [41]. By doing this, models are guaranteed to be reliable and operate as expected even if data distributions change.

A/B Testing for Model Evaluation:

By putting A/B testing approaches into practice, it is possible to compare various model configurations or versions in actual-world situations [42]. This makes it easier to evaluate models quickly and choose the best model iteration.

Continuous Model Monitoring and Re-evaluation:

Establishing ongoing monitoring pipelines with deployed models helps teams to check performance indicators and spot drift in real time [43]. Continuous Model Monitoring as well as

Journal of Technology and Systems
ISSN : 2788-6344 (Online)
Vol. 6, Issue No. 2, pp. 56 – 72, 2024
www.carijournals.org

Re-evaluation. As fresh data comes in, this feedback loop makes sure that the models are correct and current.

Model Versioning and Deployment Pipelines:

Updated models can be easily integrated into manufacturing environments with the help of versioning techniques for models that are consistent with agile development methodologies and efficient deployment pipelines [44]. By doing this, it is made sure that the most modern and efficient models are always in use.

4.3 Cross-Functional Team Dynamics in Data-Intensive Agile Environments

Effective interconnected team cooperation is essential for the successful implementation of AI/ML initiatives in heavily data-driven Agile environments [45]. The topic at hand focuses on maximizing the organization and collaboration of teams with various specialties to handle the challenges given by projects with plenty of data.

***Role Delineation and Expertise Integration**:*

By clearly outlining the roles and duties of team members, such as data scientists, technology engineers, subject matter experts, as well as Agile practitioners, it is possible to make sure that each person contributes their unique expertise to the overall project [46]. This encourages a collaborative method of problem-solving.

Continuous Communication and Knowledge Sharing:

Agile places a strong emphasis on frequent communication. This calls for periodic reviews on the quality of data, preprocessing development, and the effectiveness of models in data-intensive initiatives [47]. Team members can better understand and value one another's contributions through the sharing of knowledge as well as possibilities for cross-training.

Cross-Functional Pairing and Collaboration:

A comprehensive approach is fostered by encouraging cooperation in programming and group problem-solving sessions involving data researchers and developers [48]. More successful approaches and a common knowledge of the project's goals are produced by this lively exchange of thoughts and abilities.

Journal of Technology and Systems
ISSN : 2788-6344 (Online)
Vol. 6, Issue No. 2, pp. 56 – 72, 2024
www.carijournals.org

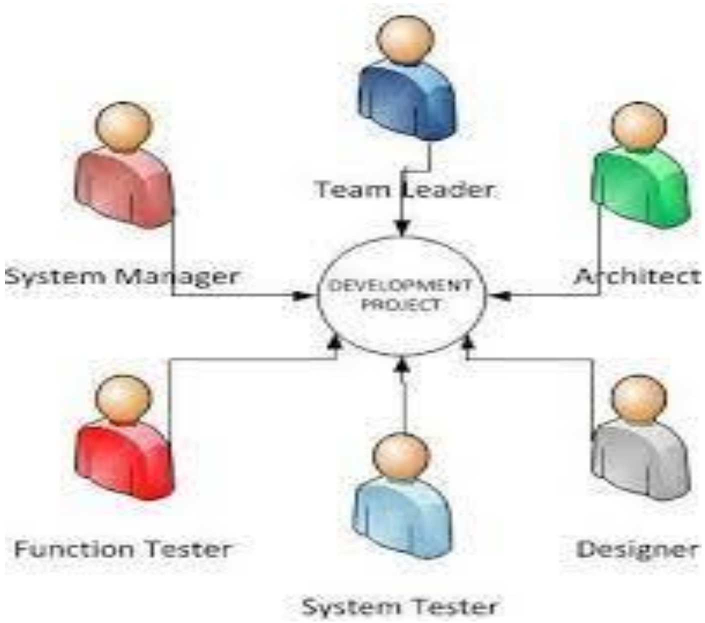

Fig. 7

Agile Ceremonies Tailored for Data-Intensive Projects:

The entire team will be in agreement with the undertaking's data-centric goals if Agile protocols like sprint scheduling, and daily stand-ups, including retrospectives are modified to include tasks involving data and conversations [49].

Feedback-Driven Iterations and Adaptations:

Regular loops of input from the group's constituents and end users, as well as from the team itself, offer insightful suggestions for improving the team's methodology [50]. Rapid response to altering data requirements is made possible by this iterative procedure.

Team Empowerment and Autonomy:

Giving teams the freedom to decide how to preprocess data, design models, and execute those models fosters accountability as well as ownership [51]. As a result, more creative and practical solutions are produced.

4.4 Tools and Technologies for Agile AI/ML Implementation

Agile approaches require a strong technology foundation that can meet the particular requirements of data-intensive projects for effective AI/ML integration. For flawless creation and operation, it is essential to use the right tools and technologies.

Containerization Platforms (e.g., Docker, Kubernetes):

Journal of Technology and Systems
ISSN : 2788-6344 (Online)
Vol. 6, Issue No. 2, pp. 56 – 72, 2024
www.carijournals.org

Agile workflows may quickly reproduce and maintain AI/ML models along with their dependencies thanks to containerization, which enables consistent deployment across various environments [52].

Continuous Integration/Continuous Deployment (CI/CD) Pipelines:

The development, testing, and deployment of code and models are all automated using CI/CD pipelines. This facilitates the quick iteration and deployment of AI/ML components during Agile sprints.

Fig. 8

Data Versioning and Version Control Systems (e.g., Git):

In projects that involve a lot of data, controlling dataset versions is essential. Data changes can be tracked using version management platforms, which are frequently used in the creation of programs and ensure reproducibility as well as traceability.

Data Pipelining and Workflow Orchestration (e.g., Apache Airflow):

Automating the creation and administration of complicated data workflows are made easier by tools including Apache Airflow [53]. They enable the coordination of tasks including model training, implementation, as well as information preprocessing inside Agile iterations.

EVALUATION AND CONCLUSION

5.1: Critical Evaluation

The combination of Agile techniques and AI/ML is a challenging but potentially fruitful project. Agile's iterative methodology fits well given the transient nature of AI/ML development, although coping with the inherent complexity of data-intensive projects presents difficulties. A thorough framework addressing the interaction between data complexity and sprint preparation is

Journal of Technology and Systems
ISSN : 2788-6344 (Online)
Vol. 6, Issue No. 2, pp. 56 – 72, 2024
www.carijournals.org

conspicuously lacking, despite the fact that existing research offers insightful analysis of tactics and tools. The technical methodology put forth uses a deductive strategy, interpretivist philosophy, as well as a descriptive design to provide a structured way to look into this integration. Recognizing potential restrictions is crucial, though. The depth of insights could be constrained by a dependence on secondary data, and certain conclusions could change due to the continuous nature of AI/ML technology. While the suggested themes address important concerns, additional elements including concerns about ethics and complying with regulations may call for more research.

5.2 Research recommendation

Future research on integrating AI/ML with agile methodologies should focus on conducting empirical studies to validate and refine the proposed technical approaches. It is important to assess the practical effectiveness of these methods through real-world trials and case studies. Additionally, ethical and regulatory aspects of AI/ML projects must be explored to ensure that agile teams can handle sensitive data without compromising the efficiency and adaptability of their processes. Promoting interdisciplinary collaboration and investing in training and skill development are crucial for equipping teams with the necessary expertise in agile and AI/ML practices. Organizations should also stay updated with technological advancements and establish feedback loops with stakeholders to adapt integration strategies based on practical feedback and evolving requirements. Lastly, documenting best practices and lessons learned from implementing these strategies will help in refining and replicating successful approaches across different contexts.

5.3 Future work

Future research in this field must concentrate on emerging technology and changing business practices as it delves deeper into the convergence of AI/ML and Agile techniques. It would be beneficial to conduct long-term studies to monitor the development of integration techniques and their effects on project outcomes. Future research must also focus on how to incorporate ethical AI concepts and responsible AI practices into Agile frameworks [55]. This subject would be advanced by examining the scalability of the suggested approaches for massive amounts of data-intensive projects and assessing their applicability in various industry domains. Finally, it would be a worthwhile direction for future studies to perform empirical studies to confirm the success of the suggested solutions in real-world circumstances.

5.4 Conclusion

In conclusion, the research effectively illustrates the nuanced interplay between agile methodologies and AI/ML project management, particularly in the context of sprint planning and data complexity. The study's recommendations provide a systematic framework that not only addresses the unique challenges posed by integrating agile practices with AI/ML development but also enhances the efficiency and effectiveness of such projects. The adoption of cross-functional teams, use of advanced technological tools, and emphasis on continuous validation are critical to navigating the complexities of data-driven environments. This research significantly contributes to the theoretical and practical understanding of agile and AI/ML integration, offering a roadmap for organizations striving to optimize their project outcomes in this innovative and rapidly evolving field.

REFERENCES

[1] Mattsson, U., 2023. *Controlling Privacy and the Use of Data Assets-Volume 2: What is the New World Currency–Data or Trust?*. CRC Press.

[2] Thiée, L.W., 2021. A systematic literature review of machine learning canvases. *INFORMATIK 2021*.

[3] Mapetu, J.P.B., Chen, Z. and Kong, L., 2019. Low-time complexity and low-cost binary particle swarm optimization algorithm for task scheduling and load balancing in cloud computing. *Applied Intelligence*, *49*, pp.3308-3330.

[4] El Emam, K., Mosquera, L. and Hoptroff, R., 2020. *Practical synthetic data generation: balancing privacy and the broad availability of data*. O'Reilly Media.

[5] RM, S.P., Bhattacharya, S., Maddikunta, P.K.R., Somayaji, S.R.K., Lakshmanna, K., Kaluri, R., Hussien, A. and Gadekallu, T.R., 2020. Load balancing of energy cloud using wind driven and firefly algorithms in internet of everything. *Journal of parallel and distributed computing*, *142*, pp.16-26.

[6] Ghasedi, K., Wang, X., Deng, C. and Huang, H., 2019. Balanced self-paced learning for generative adversarial clustering network. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 4391-4400).

[7] Ghasedi, K., Wang, X., Deng, C. and Huang, H., 2019. Balanced self-paced learning for generative adversarial clustering network. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 4391-4400).

[8] Xu, Z., Shen, D., Nie, T. and Kou, Y., 2020. A hybrid sampling algorithm combining M-SMOTE and ENN based on Random forest for medical imbalanced data. *Journal of Biomedical Informatics*, *107*, p.103465.

[9] Shan, W., Qiao, Z., Heidari, A.A., Chen, H., Turabieh, H. and Teng, Y., 2021. Double adaptive weights for stabilization of moth flame optimizer: Balance analysis, engineering cases, and medical diagnosis. *Knowledge-Based Systems*, *214*, p.106728.

[10] Deng, W., Xu, J., Song, Y. and Zhao, H., 2021. Differential evolution algorithm with wavelet basis function and optimal mutation strategy for complex optimization problem. *Applied Soft Computing*, *100*, p.106724.

[11] Fischer, D., Brettel, M. and Mauer, R., 2020. The three dimensions of sustainability: A delicate balancing act for entrepreneurs made more complex by stakeholder expectations. *Journal of Business Ethics*, *163*, pp.87-106.

[12] Song, S., Wang, P., Heidari, A.A., Wang, M., Zhao, X., Chen, H., He, W. and Xu, S., 2021. Dimension decided Harris hawks optimization with Gaussian mutation: Balance analysis and diversity patterns. *Knowledge-Based Systems*, *215*, p.106425.

[13] Van Chien, T., Björnson, E. and Larsson, E.G., 2020. Joint power allocation and load balancing optimization for energy-efficient cell-free massive MIMO networks. *IEEE Transactions on Wireless Communications*, *19*(10), pp.6798-6812.

[14] Barcelo, A., Queralt, A. and Cortes, T., 2023. Enhancing iteration performance on distributed task-based workflows. *Future Generation Computer Systems*, *149*, pp.359-375.

[15] Li, G., Shi, L., Chen, Y., Chi, Y. and Wei, Y., 2022. Settling the sample complexity of model-based offline reinforcement learning. *arXiv preprint arXiv:2204.05275*.

[16] Chen, Y., Ning, Y., Slawski, M. and Rangwala, H., 2020, December. Asynchronous online federated learning for edge devices with non-iid data. In *2020 IEEE International Conference on Big Data (Big Data)* (pp. 15-24). IEEE.

[17] Hu, J., Chen, H., Heidari, A.A., Wang, M., Zhang, X., Chen, Y. and Pan, Z., 2021. Orthogonal learning covariance matrix for defects of grey wolf optimizer: Insights, balance, diversity, and feature selection. *Knowledge-Based Systems*, *213*, p.106684.

[18] Neelakandan, S. and Paulraj, D., 2021. An automated exploring and learning model for data prediction using balanced CA-SVM. *Journal of Ambient Intelligence and Humanized Computing*, *12*, pp.4979-4990.

[19] Zhao, S., Song, J. and Ermon, S., 2019, July. Infovae: Balancing learning and inference in variational autoencoders. In *Proceedings of the aaai conference on artificial intelligence* (Vol. 33, No. 01, pp. 5885-5892).

[20] Dimakopoulou, M., Zhou, Z., Athey, S. and Imbens, G., 2019, July. Balanced linear contextual bandits. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 33, No. 01, pp. 3445-3453).

[21] El-Kenawy, E.S. and Eid, M., 2020. Hybrid gray wolf and particle swarm optimization for feature selection. *Int. J. Innov. Comput. Inf. Control*, *16*(3), pp.831-844.

[22] Nicholls, H.L., John, C.R., Watson, D.S., Munroe, P.B., Barnes, M.R. and Cabrera, C.P., 2020. Reaching the end-game for GWAS: machine learning approaches for the prioritization of complex disease loci. *Frontiers in genetics*, *11*, p.350.

[23] Ezugwu, A.E., Ikotun, A.M., Oyelade, O.O., Abualigah, L., Agushaka, J.O., Eke, C.I. and Akinyelu, A.A., 2022. A comprehensive survey of clustering algorithms: State-of-the-art machine

learning applications, taxonomy, challenges, and future research prospects. *Engineering Applications of Artificial Intelligence*, *110*, p.104743.

[24] Yang, S., He, R., Zhang, Z., Cao, Y., Gao, X. and Liu, X., 2020. CHAIN: cyber hierarchy and interactional network enabling digital solution for battery full-lifespan management. *Matter*, *3*(1), pp.27-41.

[25] Azizjon, M., Jumabek, A. and Kim, W., 2020, February. 1D CNN based network intrusion detection with normalization on imbalanced data. In *2020 international conference on artificial intelligence in information and communication (ICAIIC)* (pp. 218-224). IEEE.

[26] de Bello, F., Botta‐Dukát, Z., Lepš, J. and Fibich, P., 2021. Towards a more balanced combination of multiple traits when computing functional differences between species. *Methods in Ecology and Evolution*, *12*(3), pp.443-448.

[27] Wang, M., Zhang, Q., Lam, S., Cai, J. and Yang, R., 2020. A review on application of deep learning algorithms in external beam radiotherapy automated treatment planning. *Frontiers in oncology*, *10*, p.580919.

[28] Yang, L., Yao, H., Wang, J., Jiang, C., Benslimane, A. and Liu, Y., 2020. Multi-UAV-enabled load-balance mobile-edge computing for IoT networks. *IEEE Internet of Things Journal*, *7*(8), pp.6898-6908.

[29] Seyyedabbasi, A. and Kiani, F., 2023. Sand Cat swarm optimization: A nature-inspired algorithm to solve global optimization problems. *Engineering with Computers*, *39*(4), pp.2627-2651.

[30] Cai, Y., Luan, T., Gao, H., Wang, H., Chen, L., Li, Y., Sotelo, M.A. and Li, Z., 2021. YOLOv4-5D: An effective and efficient object detector for autonomous driving. *IEEE Transactions on Instrumentation and Measurement*, *70*, pp.1-13.

[31] Küstner, T., Fuin, N., Hammernik, K., Bustin, A., Qi, H., Hajhosseiny, R., Masci, P.G., Neji, R., Rueckert, D., Botnar, R.M. and Prieto, C., 2020. CINENet: deep learning-based 3D cardiac CINE MRI reconstruction with multi-coil complex-valued 4D spatio-temporal convolutions. *Scientific reports*, *10*(1), p.13710.

[32] Bratianu, C., 2020. Toward understanding the complexity of the COVID-19 crisis: A grounded theory approach. *Management & Marketing*, *15*.

[33] Li, X., Wang, K., Lyu, Y., Pan, H., Zhang, J., Stambolian, D., Susztak, K., Reilly, M.P., Hu, G. and Li, M., 2020. Deep learning enables accurate clustering with batch effect removal in single-cell RNA-seq analysis. *Nature communications*, *11*(1), p.2338.

[34] Zhang, J., Guo, H., Liu, J. and Zhang, Y., 2019. Task offloading in vehicular edge computing networks: A load-balancing solution. *IEEE Transactions on Vehicular Technology*, *69*(2), pp.2092-2104.

[35] Giatrakos, N., Alevizos, E., Artikis, A., Deligiannakis, A. and Garofalakis, M., 2020. Complex event recognition in the big data era: a survey. *The VLDB Journal*, *29*, pp.313-352.

[36] Johnson, J.M. and Khoshgoftaar, T.M., 2019. Survey on deep learning with class imbalance. *Journal of Big Data*, *6*(1), pp.1-54.

[37] Kasai, J., Pappas, N., Peng, H., Cross, J. and Smith, N.A., 2020. Deep encoder, shallow decoder: Reevaluating non-autoregressive machine translation. *arXiv preprint arXiv:2006.10369*.

[38] Lipare, A., Edla, D.R. and Kuppili, V., 2019. Energy efficient load balancing approach for avoiding energy hole problem in WSN using Grey Wolf Optimizer with novel fitness function. *Applied Soft Computing*, *84*, p.105706.

[39] Ayre, J. and McCaffery, K.J., 2022. Research Note: Thematic analysis in qualitative research. *J Physiother*.