## **Point-by-Point Response to Reviewers**

Date: 03-Jan-2025

Manuscript: **An Empirical Study of the Naive REINFORCE Algorithm for Predictive Maintenance**

Dear Editor-in-Chief,

Discover Applied Sciences

Thank you for providing us an opportunity to revise and submit our manuscript, “An Empirical Study of the Naïve REINFORCE Algorithm for Predictive Maintenance”.

We thank the reviewers for their comments. We have addressed all the comments and it has greatly helped enhance our manuscript and we have formally thanked and acknowledged them in our manuscript.

In this point-by-point response document we address all the observations and provide page and section references so that the revisions can be traced back to the revised manuscript. In the revised manuscript the changes are highlighted in light-blue.

**Summary**:

* Assistant Editor: Article Highlights were missing. Added to manuscript.
* Reviewer # 1: 6 comments (9 total observations), addressed on: Pages 2 to 6
* Reviewer # 2: 9 comments (12 total observations), addressed on: Pages 2 to 6

Kind regards,

Dr. Satish Kumar

Symbiosis International (Deemed University)

**Reviewer #1 Comments**

We **thank** the Respected Reviewer for the positive and encouraging comments, along with the extremely meticulous review of our manuscript. The Reviewer suggested adding hyperparameter analysis that improved the depth of our empirical study and is especially pertinent to the field of reinforcement learning (RL). We now cover these, and this has helped improved the technical content of our manuscript.

|  |  |
| --- | --- |
| **Comment 1** | While the study’s focus on “untuned” RL models is its core premise, exploring minor variations in key hyperparameters could provide richer insights. For instance, a broader evaluation of learning rates for REINFORCE may clarify whether its exceptional performance stems from unique dynamics or optimal settings. |
| Response 1 | Thank you for this very significant suggestion. This has really helped enrich the manuscript technically. Your suggestion on analyzing the impact on variations of the hyperparameters on REINFORCE added to the novelty of our research.  We first studied existing research to identify which hyperparameters to analyze. We analyzed learning-rate, the discount factor γ and network activation functions (Tanh and the ReLU). We then designed and ran experiments to analyzed the results. Interesting insights were generated and discussed the findings in the Discussion Section.  Addressed on Pages 29, 30 and 31 in section “Sensitivity analysis of hyper parameters” and the Discussion Section on page 31, 32 and 33. We added Figures 23, 24, 25 and 26. |
| **Comment 2** | REINFORCE’s training time is reported as significantly longer than the other algorithms, with considerable variance. However, there is no discussion of strategies to reduce its training time. I suggest adding a cost-benefit analysis comparing training time to achieve precision or recall. |
| Response 2 | Thank you for this suggestion. Your previous suggestion on analyzing hyperparameters, helped us unveil some of these strategies. This also helped us review some established research on this topic, specifically François-Lavet (2016).  We mention strategies around the discounting rate, the learning rate and the use of the ReLU activation layer.  Addressed on Pages 30, 31 and 33. Section “Sensitivity analysis of hyper parameters”, “Impact of hyperparameter setting on training time” and “Compute cost of performance improvement”. Fig. 25 and Fig. 26.  Francois-Lavet, V., Fonteneau, R., Ernst, D.: *How To Discount Deep Reinforcement Learning: Towards New Dynamic Strategies*. CoRR abs/1512.02011 (2016) |
| **Comment 3** | The REINFORCE algorithm reported pool recall, especially in complex multivariate states' experiments. The recall is critical in industrial maintenance tasks because failing to replace a tool at the right time can lead to catastrophic failure or significant downtime. Therefore, this poses a risk to reliability in high-stakes environments. |
| Response 3 | We agree. Our focus was on precision which is driven by lower false-positives (FPs) i.e. reducing unnecessary replacements. However, we agree, in general industrial scenarios, the cost of tool is much “lower” compared to cost-of-quality.  We have therefore modified our text and state the importance of considering the metric on the basis of the industrial application and considering F1 as a more balanced measure. We provide suggestions on using suitable value for F-beta (β = 2.0) when importance to recall is desired. The abstract is also modified to include recall metrics.  We modified the text to state that “for this research we consider β = 0.5”. This is probably applicable in situations where the actual process of replacing part itself is expensive and a considered a hazard for the maintenance staff, for example maybe wind-turbine maintenance.  Addressed on Page 20: Evaluation metrics. |
| **Comment 4** | While the concept of testing untuned models makes the results more accessible to practitioners, it also limits the study’s depth. Practically tuned models are preferred for deployment. Also, comparing untuned versions ignores the fact that other algorithms (e.g., PPO or DQN) might outperform REINFORCE when tuned. |
| Response 4 | This is correct and is a point we had missed stating. We mention that advanced algorithms provide a richer set of hyperparameters and that the tuned versions of these algorithms will perform better than the naıve REINFORCE. However, we admit that a deeper study by tuning these algorithms is a limitation of our scope and is an important subject for future research.  Addressed on Page 34, Discussion Section under “Limitations and future scope for research”. |
| **Comment 5** | The paper attributes REINFORCE’s surprising performance to factors like activation functions (ReLU vs. Tanh) and learning rate. Still, there is no discussion of architectural simplicity as a driver for better performance in smaller data environments. I suggest including baseline comparisons against simpler supervised learning models to contextualize the value of RL. |
| Response 5 | Thank you for highlighting this. We analyze the view of established researchers on architectural aspects of RL implementations as well as REINFORCE’s architectural simplicity as a possible reason for its better performance.  We surveyed **12 articles** that have used supervised machine learning (ML) to solve a similar predictive maintenance (PdM) problem. In Table 1 we identify the ML technique used, the type of time-series problem (univariate or multi-feature) it was addressing and the actual PdM use-case it solves. In the Discussion section, we briefly compare the two methods, RL and ML methods.  Addressed on Page 7, Table 1 and Page 33, Discussion Section. |
| **Comment 6** | Although statistical metrics and tests are provided, the paper does not offer confidence intervals for most metrics, limiting interpretation of results’ robustness. Also, variability in performance across datasets is not well analyzed or explained. |
| Response 6 | We have now added 95% confidence intervals to all the results tables in the main text. The new plots from the Sensitivity Analysis contain 95% confidence intervals. All the performance plots already had error bars.  Revised tables in the Results Section: Table 6 pg. 21, Table 7 pg. 23, Table 8 pg. 24, Table 9 pg. 25 and Table 10, pg 27. Sensitivity Analysis section Figures 23 and 24 on page 30.  The variability in performance across datasets is a result of the high variation in features. We have now explanation on in the ‘Actual tool wear data’ sub-section of the ‘Implementation details’ section and added plots to show this variation, see Fig. 6, page 12. |

----------- End of “Reviewer #1 Point-by-point” section -----------

**Reviewer #2 Comments**

We would like to sincerely **thank** the Respected Reviewer for the positive comments and appreciate the additional related areas highlighted, that will add value to the research community. We also thank the Reviewer for pertinent observations; addressing these has assisted in adding new content and improving existing sections. Thank you.

|  |  |
| --- | --- |
| **Comment 1** | Please include the stats of the paper found and also the methodology employed in this paper. |
| Response 1 | xxx.  Robust RL: xx  Risk-sensitive RL: xx  We covered a few other related fields in the original manuscript while covering some as part of this revision.  In Section 12, Page 63, under the head “Inherent challenges of implementing deep RL” we covered (1) Stability of RL algorithms (2) Reward shaping and sensitivity to scale (3) Over-fitting in model-free RL (4) Hyperparameter tuning and (5) Effect of adding a noise parameter to enable exploration and generalize learning in the case of continuous action spaces. In Section 13.4 (page 71), we touched upon (6) reward shaping and reward learning through Inverse RL and (7) Curriculum learning on the same page.  Thanks to Reviewers, we now cover the following four fields: In Section 13, page 70, we covered Meta-RL under “RL methods for enabling agents to handle”. Page 70, Section 13.2 covers Batch- or Offline- RL fields and is connected to its application “complex systems made up of multiple sub-systems” on page 71. On page 72, Section 13.6, we cover Hierarchical RL and the options framework. Finally, on page 73, Section 13.7, we cover maintenance of controllers.  Overall, this has helped us now cover about 11 related fields, and this has really enriched the manuscript. |
| **Comment 2** | There is no discussion on the cost effectiveness of the method. What is the computational complexity? What is the runtime? Please include such discussions. |
| Response 2 | Thank you for highlighting this very important observation. While we have invested significant efforts to cover technical details of research articles, throughout this manuscript; the details remain *scattered* over the entire manuscript and hence the section titled “Algorithms applied by researchers (RQ-4)” does indeed appear light.  We have done two things to address this:   1. Table 8, pages 57 to 62: As suggested by the Reviewer, we have summarized *select* articles in the format suggested and added a deeper level of technical understanding of the techniques applied (instead of just high-level root algorithm). This is a new table titled “Algorithm details: PdM scenarios and challenges addressed by researchers” 2. We have now added a note to help the reader find the technical analysis of articles that are spread across our manuscript. Together the above table, we now cover about **50 articles**, in some form of details.   We repeat the note we added here, along with *page numbers* to assist in reviewing.  “Algorithmic and technical details of a number of articles were studied in the previous sections. Section 9.2 *(page 39) was* devoted to the SMART family of algorithms designed specifically for solving the PdM problem. Detailed case-studies in Section 9.3 *(page 41)*;followed by Section 10 *(page 45)* with an industry-oriented study; covering rotating and milling machines, hydraulic actuators, joint optimization of industrial problems. Assembly lines and multi-machine systems were covered, and Table 7 *(page 54)* summarized implementation details across multiple articles. Finally, Table 8 *(page 57)* summarizes select research articles by listing the PdM scenario they address, research challenges they tackle and the algorithmic techniques used to address these challenges.” |
| **Comment 3** | To have an unbiased view in the paper, there should be some discussions on the limitations of the method. |
| Response 3 | We agree that there is a significant xx |
| **Comment 4** | Neither the novelty nor the uniqueness of the research is established. |
| Response 4 | We agree that there is a significant xx |
| **Comment 5** | Authors need to add more latest references from the years 2022 and 2024. |
| Response 5 | We agree that there is a significant xx |
| **Comment 6** | Abstract needs to relook and highlight the scope and then add what is the aim/Objective of the paper, also highlight the numerical Findings and compared to existing works to justify that the training set model works better and what is the overall analysis. |
| Response 6 | We agree that there is a significant xx |
| **Comment 7** | Suggested to relook the conclusion section and highlight the open issue for further research contribution. The quality of the figures and tables need to be checked. |
| Response 7 | We agree that there is a significant xx |
| **Comment 8** | The technical contribution of this research is not adequately described in the abstract. I advise rewriting it. |
| Response 8 | We agree that there is a significant xx |
| **Comment 9** | The methods part is poorly designed and needs improvement to include more evidence on the adequacy of the research procedure. |
| Response 9 | We agree that there is a significant xx |
|  |  |
|  |  |

----------- End of “Reviewer #2 Point-by-point” section -----------