Improving the Efficiency of Work Order Management by Infusing AI-Empowered Automation

Roy Abitbol, [1] **Eyal Cohen,** [1] **Shlomit Gur,** [1] **Muhammad Kanaan,** [1] **Lior Limonad,** [1] **Hadar Mulian** [1]

[1] IBM Research - Haifa, Israel

Abstract

Numerous efforts in the field of AI are aimed at automating repetitive tasks in business processes. In some cases, processes are reliant upon human decision-making. We hypothesize that providing AI-generated recommendations at key decision points in a business process assists in reducing the workload and errors. We implemented two predictive AI models, providing recommendations in two key stages of a work order management process. We tested the models performance for accuracy in a case study and received positive feedback from test users. In our current work (in progress), we evaluate the effectiveness of the AI recommendations using an experiment, recording the performance metrics with and without the help of the models. We aim to test for the statistical significance of the assistance effect with respect to a series of variables.

Introduction

Recent advancement in the intersection between Artificial Intelligence (AI) and Business Process Management (BPM) focus on automating repetitive tasks (Chakraborti et al. 2020). Other repetitive tasks require a high degree of critical thinking and cognitive awareness and as a result they mandate a decision carried out by human agents (Lacity and Willcocks 2016). On these tasks, as well, AI has the potential to improve by aiding humans in their decision making through AI-generated recommendations (Williams and Olajide 2022). In our work, we focus on these AI-aided automation tasks, proposing to improve employee performance rather than replacing them entirely.

One prototypical example of a Business Process (BP) comes from the domain of Work Order (WO) management where a systematic approach is used to manage the maintenance life cycle. In this work, we focus on corrective WOs, managed on IBM Maximo. The information captured in the WO management process contains the description of the tasks, the labor, the materials, the services, and the tools needed to complete each task.

A typical corrective WO management process consists of a set of stages in its life cycle, usually: creation, approval, prioritization, assignment, scheduling, execution, documentation, closure, and analysis (Figure 1) (Alayna Giesting 2020). WO processes may vary in their sequence of stages as well as allow for loops and other flow variations, in different companies or different types of work. The tasks at most of these stages, require a certain human decision in order to advance the WO to the next stage. For example, during the approval stage, a maintenance manager (or coordinator) decides whether the WO should be approved or not.

Internal user studies show that despite the apparent simplicity of some of those decisions, they still impose work load (mostly mental) on the decision maker and they incur a certain cost, primarily in the form of time spent by the decision makers. Furthermore, possible decision mistakes made by humans in the earlier stages of the process might result in greater costs downstream, because of later corrections having to be made. Taking into consideration the numerous decisions in these processes and the countless instances of these processes (i.e. WOs), the cost incurred by these decisions may be rather high. Our goal is to show that offering AI-based decision support for key decisions will improve the efficiency of the process by reducing the workload and errors.

Method

We focus on two key decisions, contributing to the burden of workload for maintenance managers: *WO approval* and *WO assignment* (Figure 1, dark blue rectangles). The former relates to the action of approving a new WO request, by reviewing its information and determining if it describes a valid and legitimate need. The latter, describes the action of assigning (or dispatching) the WO to the most suited assignee (i.e. person), based on the skills and crafts needed to resolve it as well as the availability.

To address these decision points, we developed two AI models aimed at assisting humans in making the decisions in these stages by providing recommendations for suitable actions. We acquired a set of WO records and used them to train two predictive models. After training, the models are capable of reading a WO's attributes and producing a prediction whether a WO should: a) be approved or not, and b) who should it be assigned to. The dataset used contained roughly one million logs of WOs from an organization, describing the work related to maintenance and handling at numerous facilities world-wide.

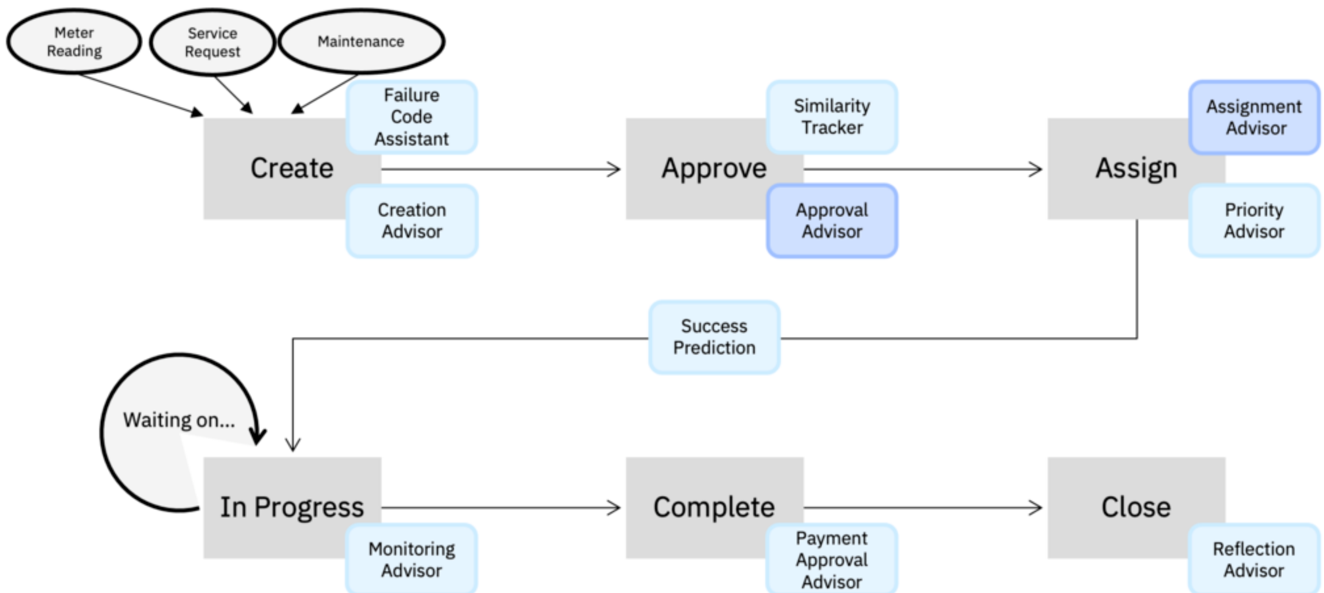

Figure 1: A possible flow of the key stages in the WO management BP. Each grey box presents a stage in the process. Examples for possible triggers for WO creation are indicated as a grey ovals at the top left, where some are automatic and some manual. In each stage of the WO BP there are opportunities for AI-based recommenders, marked in blue. We focus on two recommenders, marked in darker blue, assisting with the approval of WOs and with the assignment (i.e. dispatching) of WOs.

Results and Conclusions

After several iterations for training and testing the model with tree-based ensembles, the best model (XGBoost, (Chen and Guestrin 2016)) reached a Macro F1-score of 0.70. It is worth noting that the data used for the approval recommendation task is quite imbalanced with strong tendency towards an approval of WOs (roughly 97% approved), hence, the accuracy score for identifying the "approved" correctly was over 98% and for the "cancelled" over 50%.

The assignment model was also based on a tree-based ensemble model (XGBoost) but unlike the approval which had a binary class, the assignment model had to deal with multiple classes representing the list of persons who could be assigned for the task. Since the number of unique persons, eligible for tasks world-wide is quite large, we trained and ran the model on a per-site basis. We tested the model on two specific sites. For site A, the Weighted F1-score for a correct prediction of the assignee based on the single top-most prediction was 0.62 and when testing if the correct prediction is within the top-3 predictions the Weighted F1-score increased to 0.82. For site B, the Weighted F1-score was 0.74, and similarly to site A, when tested the top-3 prediction, the score improved to 0.89.

The models' recommendations were embedded in a mock user interface, mimicking a real production environment where users are performing their daily approval and assignment routine. Four users (maintenance managers and coordinators), from four different sites, were presented with the recommendations on the mock environment before actually carrying out the operations on the actual environment. We have also asked the users of the system to provide feedback for the model using a "thumbs" up or down for each record, indicating their agreement with the recommendation. Furthermore, at the end of the trial period, users were asked to provide a verbal statement with their impressions from the model. The results of the users' "thumbs" feedback were over 90% positive (i.e. "up"), namely they accepted the models' recommendation in more than 90% of the cases. Furthermore, all the verbal statements that were given were positive. To quote one such statement: *"This tool, when it is incorporated to Maximo, will be of great help, because today we have to do these analyzes in a manual way"*. This encouraging feedback lays the foundation for the next stage in this work in which we measure the effectiveness of the AI recommenders on the process.

Work in progress

In order to establish that these attempts at Intelligent Automation are in fact beneficial in improving the performance of the maintenance team, we are currently conducting an experiment. Our hypothesis is that with the help of AI recommendations the human decision making is easier and therefore quicker and less error prone. For this purpose we measure the duration of these two decision making points and the rate of errors performed by the Maintenance managers, with and without an AI recommendation. Furthermore, we measure the impact of presenting the users with a "recommendation confidence", in order to establish its effectiveness. Finally, the hypothesis analysis will be tested for statistical significance of the assistance effect with respect to a series of dependent variables such as perception of usefulness, etc.

References

Alayna Giesting. 2020. How to Create the Best Work Order Process: Ensuring Maintenance Success. https://www.gofmx.com/blog/work-order-process/. Accessed: 2022-11-27.

Chakraborti, T.; Isahagian, V.; Khalaf, R.; Khazaeni, Y.; Muthusamy, V.; Rizk, Y.; and Unuvar, M. 2020. From robotic process automation to intelligent process automation. In *International Conference on Business Process Management*, 215–228. Springer.

Chen, T.; and Guestrin, C. 2016. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, 785–794.

Lacity, M. C.; and Willcocks, L. P. 2016. A new approach to automating services. *MIT Sloan Management Review*, 58(1): 41–49.

Williams, O. C.; and Olajide, F. 2022. Towards the Design of an Intelligent Automation Framework for Business Processes. In *2022 5th International Conference on Information and Computer Technologies (ICICT)*, 13–17. IEEE.