Infusing Work Order Management with AI

Roy Abitbol, Eyal Cohen, Shlomit Gur, Muhammad Kanaan, Lior Limonad, Hadar Mulian

Goal

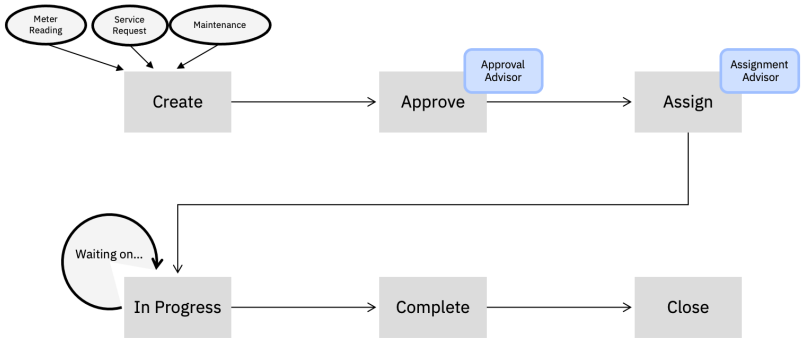
Improve employee performance in Work Order (WO) Management by automating suggestions for WO approval & assignment using AI

Hypotheses

- Assistance by AI models improves employee **performance** in terms of **correctness** and **duration**
- Presenting AI model's **confidence** level influences the performance

Work Order Lifecycle

- Typical WO process and the suggested AI models



Data

- Maintenance WO records from a global resource organization
- +1M records globally, with +20K records for the case study site
- Approval is a binary decision having approval rates of ~97%
- Assignment is a multi-class decision having ~30 classes per site
- On-site data is sparse & incomplete and therefore hard to predict
- Ground truth subset taken from expert actions was used as reference for all performance measurements

Models

- XGBoost based models deployed in 5 sites and tested on-site for 2 months, by a total of 10 professional users
- Results of the deployed model (F1-score) for on-site data:
- Assignment: 0.84 top-1 and 0.92 top-3, Approval: 0.71
- Models received +90% positive feedback for advices
- User testimonials: *"...this tool will be of great help..."*

Experiment

Evaluated the effectiveness of the AI assistance by measuring performance of employees in various conditions

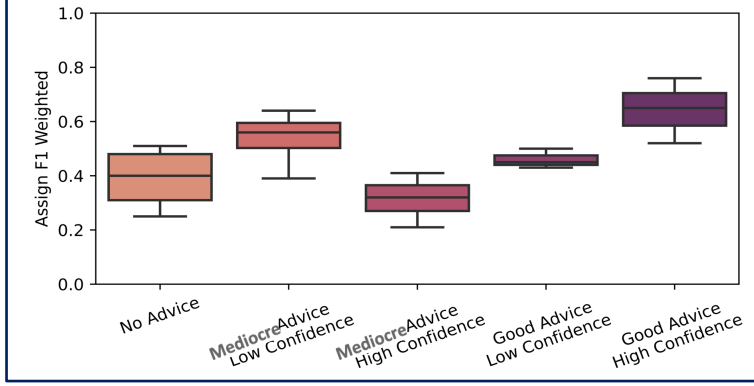
- Performance is measured by **correctness** and **duration**

Manipulated factors -

Professional users perform *Approval* and *Assignment*, first without any AI advice and then with manipulations of

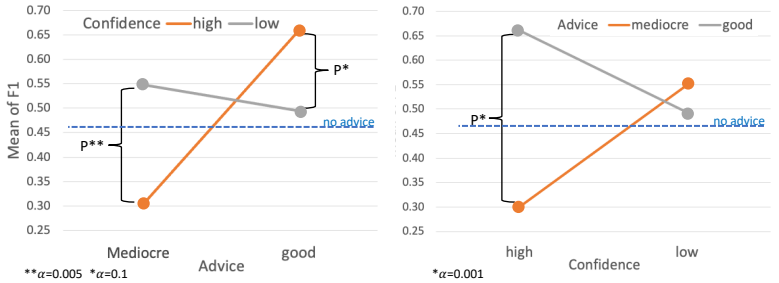
- *quality of advice*: *good* or *mediocre* advice
- confidence: *high* or *low*

Results



Results (cont.)

- Testing using 2-way ANOVA shows statistical significance of the performance for the interaction of quality and confidence (p=0.001)
- Post-hoc analysis shows significance when comparing *low* vs. *high* confidence for a *mediocre* advice (p=0.002) but only mild significance for different confidence having a *good* advice (p=0.09)



- When the confidence is *high*, a *good* advice is statistically significant compared against a *mediocre* advice (p<0.001) or against *no-advice* (p=0.02), but not so when the confidence is *low*

Conclusions and Future Work

AI recommendations and their confidence scores have significant influence on employees' performance in a WO business process:

- **When the confidence score is high, a good advice is effective compared to no-advice, but a mediocre advice might be harmful**
- **When the confidence score is low, the effect of the advice is insignificant, compared with no-advice, regardless of its quality**

In our future work we broaden the experiment:

- Additional subjects and better quantification of the duration
- Measure the subjects' perception of their performance