Seminar- scientific methods in information systems

Predictive Process Monitoring Methods: Which One Suits Me Best?

Shuaiwei YU[1]

Technical University of Munich
`shuaiwei.yu@tum.de`

Abstract. The abstract should briefly summarize the contents of the paper in 150–250 words.

Keywords: First keyword · Second keyword · Another keyword.

1 Introduction

2 Systematic Literature Review Protocol

The research questions are formulated in the systematic literature review protocol, the search protocol is defined, and the selection criteria are developed. The selection criteria are based on the selection criterion of [1] and are modified following the guideline provided in Kitchenham's article [2]. The work consisted of several stages. In the first stage, The review protocol was designed. Specifically, the research questions to be investigated were formulated, and I chose the electronic database that will be used to run the search. The selection criteria, as well as inclusion criteria, were also defined. In the next stage, the horizontal search was run. The documents acquired were then filtered according to selection criteria; in the end, the final list of documents will be generated.

The main research question (**RQ**):"*What is the development and trend of the predictive process monitoring techniques over the research period?*" can be decomposed in four sub-question as follows. The first two sub-questions are so developed in order to research the current trends and algorithmic distribution of the predictive process monitoring. **RQ1**: "*What is the distribution of the literature about predictive process monitoring over the search period?*" and **RQ2**: "*What is the distribution of the algorithms used in predictive process monitoring over the search period?*". The next sub-question focuses on the aspects of outcome that can be predicted, **RQ3**:"*What aspects can be predicted by the predictive process monitoring?*". In the last question, we want to research the reason that leads to the current trend of predictive process monitoring, i.e., **RQ4**: "What leads to the trend of predictive process monitoring?"

In the first research question, we want to visualize the distribution of the article published in order to analyze the current trend of predictive process

monitoring. With this visualization, we are able to take a deeper look into the development of interest in the industry of predictive process monitoring. The second research question examines the distribution of algorithms used for predictive process monitoring. The motivation for this is that we want to reveal which algorithms are more popular and widely used. With the first two research questions, we can derive the current research trend of predictive process monitoring. We can also analyze the possible research gaps where further academic contributions could be made. In [1], several aspects were already listed. However, the third research questions want to research whether there are new predictive aspects developed or how the new predictive process monitoring model predicts these aspects differently. The last research question tries to look into the reason that led to the current development of predictive process monitoring so that we might predict what the research direction of this field might be.

The search string must be defined so that we can query documents in an electronic database. We followed the guidelines and suggestions given by [2] and based on the work of [1], we first chose the keywords "predictive", "prediction", "business process", and "process mining". Out of these four keywords, a search string can be built: ("predictive" OR "prediction") + ("business process" OR "process mining"). Furthermore, in order to explore as many results as possible, we also decided to add the keyword "process monitoring" to acquire more specific documents, so another search string is: ("predictive" OR "prediction") + "process monitoring".

After defining the search string, we could choose electronic databases for searching. We decided to use Google Scholar as electronic databases to be used, for its completeness and simple operability. The search period was between year 2018 and 2022. We identified 263 hits initially with our keywords; after the search, we exported the results for further filtering.

The first step of the filtering is about identifying the duplicates. Since several Databases are used, and the keywords can potentially overlap, our goal is to identify the literature that appeared more than once and whose name and author are identical. Only one of the papers will be preserved for future processes. After the duplicates filtering was performed, there were in total of 153 papers identified. In the next step, we look into the title of the papers gathered. In this step, (i) the studies that are related to a different field will be removed (e.g., "A Chemical Monitoring and Prediction System in Semiconductor Manufacturing Process Using Bigdata and AI Techniques", or "Condition monitoring and prediction of solution quality during a copper electroplating process"); (ii) the documents which are not proper research papers should also be removed (e.g., "Integration of an Explainable Predictive Process Monitoring System into IBM Process Mining Suite (Extended Abstract)"). The number of papers was reduced to 114 after filtering.

Moreover, we only consider the papers published in conferences, journals, and workshops, because, in comparison to degree thesis, published articles are more professional and have a higher academic contribution. After this stage of filtering, 106 papers are left. In the next stage of filtering, we looked into the

abstract of each paper in order to evaluate its relevance. We also leave out the position paper at this stage since we only intended to look into the algorithms being used. After this filtering, 78 papers remained. In the last stage, we use the inclusion criterion *"the study should propose a novel algorithm or technique for predictive process monitoring"* to further select the papers used for our study. We selected 28 papers in total, and through a backward reference search, 4 papers were added.

Keywords	Hits	Selected	Selection criteria
allintitle:("predictive" OR "prediction")+("business process" OR "process mining")	112	19	1. Full paper published in conferences, journals or workshops, written in English 2. Is a proper research paper
allintitle:("predictive" OR "prediction")+process monitoring	151	9	3. Relate to computer science/business process/process mining area 4. Inclusion criterion: the study should propose a novel algorithm or technique for predictive process monitoring
Result horizontal search	263	28	
Add papers		4	Vertical search
Overall		32	

Fig. 1.

3 Systematic Literature Review Results

References

1. Di Francescomarino C, Ghidini C, Maggi F M, et al. Predictive process monitoring methods: Which one suits me best?[C]//International conference on business process management. Springer, Cham, 2018: 462-479.
2. Kitchenham B. Procedures for performing systematic reviews[J]. Keele, UK, Keele University, 2004, 33(2004): 1-26.