Name : SAMAROO Sharmila, Master 2 MCI-Amérique Latine

Name of your level 1 : Angélique CHAGNEAU

Keywords specific to the content : machine learning, process mining, process optimization, PN

The software used : Microsoft Excel

Source : e.g Googlescholar.com
BDA/Business_process_analysis/L2_AissataDIAW/L1_AngéliqueCHAGNEAU/L0_SAMAROOSharmila/Machine learning.pdf at main · BigData2024/BDA (github.com)

Paper title - Article : Machine learning and process mining applies to process organization :

This article was written by Edson Carvalhar Fernandes, Barry Fitzgerald, Liam Brown, Milton Borsato in 2019. This article focuses on changes in the goods and services market and the need for enterprises to become more productive and more efficient.

This article highlight the identification of :

- Opportunities
- Gaps
- Process management

Those elements are important for :

- The smart factory, machine learning and process mining as part of process optimization

Name of the process	Definition :	Main problem/ Allow to
Machine learning (ML)	Allows computer to learn without being explicitly programmed	Pr : Solution not efficient
Process mining	Allows to increase the growth in manufacturing industries	
Process optimization	Needs an efficacity of ML and PM. – the content have to be relevant, scientific and systemic analysis.	Allows to : compare result through regular meeting mapping and manual documentation

- The goal of these article is to provide different solutions to the market and factories efficiency over the last 5 years (they used articles from 2014 to *2018*) by investigating problems that are not studied in detail "*and help to determine the best research design and data collection".* (See article).

To ensure the market's effectiveness, the research methods were used: Exploratory research and PRO KNOW-C methodology. Those methods involve a number of steps.

So, they started looking for : bibliography, documentary, experimentation, case studies, books, which formed the database.

STEP 1 : The exploratory research consists of establishing search on 3 axes :

- process optimization : 17 keys words associated
- machine learning : 17 keys words associated
- Process mining : 17 keys words associated

STEP 2 : Then use “and” and “or” to associate/combinate those keywords, for example “process improvement” or “process mapping”

Once the searches where completed, they were grouped them into 5 databases :

- Academic search
- Emerald
- Insight
- IEEF
- Science direct and engineering

First, they found 3562 raw articles. Now it’s a matter of keeping only the articles that respond to their need and subject so :

STEP 3 : They used the software “Mendeley” - it allows to :

- Managing those articles
- Identifying and eliminate all duplicate articles
- Identifying published and non-published articles

At the end of this process there are 782 articles left. Now they only have articles that are interesting to their subject.

STEP 4 : they transfer these articles to Microsoft Excel and into Google Scholar filter : so they will know, which articles are the most cited - indicating the title, author name, publication date/ year.

STEP 5 : This leads to the Pareto’s principle. This means that 80% of citations concern 20% of publications. They went from 782 articles to 193 most-cited articles and 589 less-cited articles.

STEP 6 : divide the most-cited (K) articles from the least cited one (P) :

Most cited : K’s category	Least cited : P’s category
193 articles necessary to be read. Finally they kept 12 articles .	589 articles were divided into categories. Finally they kept 20 articles

STEP 7: Bibliometric analysis : *“allows to interpretation and evaluation of the bibliographic portfolio. Establish the journal’s relevance regarding is impact factor, the published year, key word, author, knowledge opportunity*.” Definition coming from the article.

- Of the 32 articles, 12 are the most cited of 2 journals =) but they are not considered “the best according to the scientific journal ranking”
- Of the 32 articles, 3 are not included

The main problem identified in this article, according to data and technologies:

- Enterprises store the" *data in database management system, but they don't have the capability to analyze and gain useful insight…the absence of adequate process planning or the development of insufficient data mapping leads to waste of time and costly resources."* Extract from the article.
- Process optimization : due to technological evolution, the proposed strategy is less effective for factories. Enterprise and factories need to improve their level of competitiveness.
- When the companies environment change, sometimes they don't achieve outcome that they desire, which can lead to uncertainty.

STEP 8 : Use PN to solve the problem of outcome – proposed objectives and resources.

PN : Resources that allow companies to achieve better outcome.

In additional AI and process management help to *"improve strategies and eliminate failure".* Extract from the article.

Using PN to :	Using predictive system ML and PM to :
"Modelling production system to support analysis"	*Predicting performance, predicts future events in a business process*
"Modelling the production planning process of a manufacture company"	*Machine Learning: detects errors and predicts failures*
"Path planning of a machine system : use by manager for easily mapping, visualizing undesired event and potential weaknesses."	*Process Mining :" is used to estimate the probabilities of human's error when required tasks are conducted ."*
"Improve the communication between machines and product in a modular production system."	

Extract from the article.

We can conclude, that this article highlights exploratory research and PRO KNOW-C methodology. These are time-consuming methodologies : to checking the documents, reading them and proposing new opportunities to ensure that market process continues to be efficient and effective. In additional, it's allowed us to understand the objectives, problems, gaps and opportunities, in process optimization by using process learning and process mining.