Resume L0 of Business Process Analysis :

FILE 1 :

- Name: Pinto Frias Adriana

- Name of your Level 1: Eliott Laine

- Paper title: Adaptive business process analysis using machine learning algorithm

- Source (e.g. scholars.google.com): scholars.google.com

- Keywords specific to the paper:

Generalized Learning Vector Quantization (GLVQ), K-Nearest Neighbors (KNN), Decision Trees, Random Forest algorithms, > NEXT(LOG), ML.LOG, Business process logs , autoencoders, Matplotlib, Graphviz, "Auto-ML", F1 score

-Al model used (e.g. Neural network, etc.) : How do they contribute the idea proposed by the paper?

Machine learning tools used :

- GLVQ, KNN, Decision Trees, and Random Forest each of them contribute to the idea proposed by the paper by their own and specific function analyzing logs and identifying patterns or rules within the data.

- Supported by a software application? (If yes, provide more details) :

- > NEXT(LOG) program : provide a framework of dataset to improve business process logs.
- ML.LOG program : facilitate the application of machine learning or analyzing and optimizing adaptive business process logs
- "Auto-ML" : automate and optimize the machine learning process used for choosing the right file.

- Summary of the main contributions (use text paragraphs, tables and if necessary, figures):

1. Introduction :

Business process analysis is an approach to identify and analyze an organization's business process, for the companies to understand how their different business areas work. However, the traditional methods of business process analysis needed a change for automated techniques like machine learning (AI). It concerns the automatic extraction of process models from event logs. Logs that are a combination of data that give information about important stages of the process.

Applying machine learning to business process analysis helps to analyze an event log of a business process that maximizes the probability of a particular result.

2. Methods :

Step 1 :

Use the > NEXT(LOG) program created by Dyllan Cartwright. This program is a custom-built software tool that provides a dataset following rules, for the machine learning models to analyze and create adaptive business process logs.

Step 2 :

Select machine learning models for analysis. Four models are used in this study:

- GLVQ (Generalized Learning Vector Quantization) : classify tasks
- KNN (K-Nearest Neighbors) : algorithm that classifies new instances based on their similarity to existing instances were K is the closest instance
- Decision Tree : makes decisions based on the features of the data.
- Random Forest : make final decision by combining multiple decision

These models were chosen based on their ability to be used with the types of data given by the program > NEXT(LOG), and their suitability for the task of identifying patterns or rules within these logs.

This training process adjusts the parameters of each model to achieve the best performance by evaluating with metrics and providing quantitative performance.

3. Data collection and implementation :

A) Tools and technology :

Tools are used to help in the decision-making processes. We can use :

Graphviz for Random Forest and Decision tree models. Graphviz is a tool for graph visualization thereby simplifying the understanding of their decision-making mechanisms.

Matplotlib for GLVQ and KNN. Is comprehensive creating static, animated, and interactive visualizations, making the decision-making processes easier.

B) Using ML.LOG :

Used to facilitate the application of machine learning to adaptive business process logs. The program works with several key stages, each stage contributes to the processes of training, evaluating, and applying machine learning models.

Training and evaluation are carried out using the ML.LOG program with the models on different datasets.

C) Home page :

Before the ML.LOG program, the user's first view is the home page, where the users can choose the model they want to use. The generated logs are saved in a CSV format, used for storing tabular data, storing the information by semicolons. The program requires data in the form of a CSV file. This data will be used as testing data. This CSV file can be generated using the > NEXT(LOG) application or can be a general CSV file. When the file is uploaded the users can proceed with their tasks.

D) Model page :

Again, users can choose different models and select the one they wish to use.

The "Auto-ML" button activates an automatically the training process that reads the contents of the "auto ml.txt" fiIe. The program helps to choose between a file generated by the > NEXT(LOG) approach or a general file.

E) Results page :

Then, the Results page will give the best model and its corresponding parameters, determined based on the F1 score that measures the performance of the classification model providing a comprehensive overview.

This section provides a comprehensive overview of the utilized parameters and various scores. Depending on the model, users can save them and once the testing is complete, comprehensive statistics will be displayed.

5. Results and analysis :

For each dataset, the results are given by the performance metrics of the machine learning models, such as F1 score, measuring the models' effectiveness in identifying patterns or rules within the adaptive business process logs.

FILE 2 :

- Name: Pinto Frias Adriana
- Name of your Level 1: Eliott Laine
- Paper title: Business process analytics and big data system
- Source (e.g. scholars.google.com): scholars.google.com
- Keywords specific to the paper:
BP life cycle (Redesign, implementation, enactment, evaluation), Process footprint, Event correlation, Process model discovery, Conformance checking and compliance monitoring, Business process enhancement, Predictive monitoring, Unstructured data, Hadoop framework, Big SQL engines —> SQL queries, General purpose system —> Apache Spark project, Big graph processing engines —> Pregel system —> Bulk Synchronous Parallel (BSP), Big stream processing engines —> Apache storm
- Al model used (e.g. Neural network, etc.) : How do they contribute the idea proposed by the paper?
Neural network : it contributes to the idea proposed by the paper by interconnecting nodes giving pattern recognition and classification. Also, helping big data systems understand the relationships from data and making predictions or decisions based on that learned knowledge leading to improved operational efficiency, agility, and competitiveness. In the text it is contributed with auto-encoders to reconstruct input data.
- Supported by a software application? (If yes, provide more details):
Apache Spark : engine for dataset analytics used for single side environment
Apache Flink : For stateful computations over unbounded and bounded data streams runs arbitrary data flow programs in parallel and in pipeline.
Big SQL engines : Making querying enterprise data easier and secure. With easy access and providing tools to manage the database system.
Pregel system : Handle massive graphs allowing developers to focus on writing vertex-centric algorithms for graph processing tasks.
Apache storm : Distribute real-time computation system and makes it easy to reliably process unbounded streams of data.
- Summary of the main contributions (use text paragraphs, tables and if necessary, figures):
I. Introduction :
Business process organization is used for a company to achieve its objectives. Now, the companies have a lot of data. Submerged by these data, enterprises would like to weave data analytic techniques into their decision making capabilities. BPM's (business process management) are large unstructured data, named big data, and support for BP (business process).
We will see the new generation of big data systems : Spark, Flink, Storm and Impala etc.
The objective is to show the link between BPM and big data technologies and provide information.

II. Process (big) data generation :
BP's (business process) lifecycle :
- (Re) design : Help an enterprise to know how and when processing users' requests by identifying and modeling collected data to design solutions to process issues.
- Configuration/ implementation : Transform BP model into a running application by targeting an execution platform with the help of IT selection and implementation.
- Enactment : BP models are put into production and created in response to triggers like receiving users' requests to see if any changes are needed. In the context of enactment, a Process Footprint can be used to analyze the impact of executing or implementing a business process on various aspects such as resources, time, costs, and outcomes.

- Evaluation : Is the post execution analysis providing prediction and suggestion models using process mining, business activity monitoring.

III. Data intensive operations in business process analytics :

Classification of BPM data-intensive operations :

- Event correlation : This represents a sub-task of constructing the event log by identifying various information for the same objects. It is used to analyze and optimize business processes by correlating events or activities occurring within the process
- Process model discovery : Automatically generate BP model by analyzing the log and providing process evaluation phase. It involves automatically extracting or inferring process models from event data recorded during the execution of business processes.
- Conformance checking and compliance monitoring : The compliance monitoring assumes a structured process execution where correlated events are generated by a process execution engine.
- Business process enhancement : Is a technique that usually relies on process-centric abstractions and query languages that enable the querying, exploration, and visualization of integrated view by analyzing the event logs.
- Predictive monitoring : This technique focuses on forecasting potential problems. It help to analyze and extract knowledge to find similarities and potential problems.

IV. Big data processing system :

Big data processing platforms :

GENERAL PURPOSE	BIG SQL SYSTEMS	BIG GRAPH PROCESSING	BIG STREAM PROCESSING
Hadoop	Spark SQL	Graph	Storm
Spark	Impala	GraphLab	S4
Flink	Presto	Graphix	Flink

V. Where can process management and big data systems meet up ?

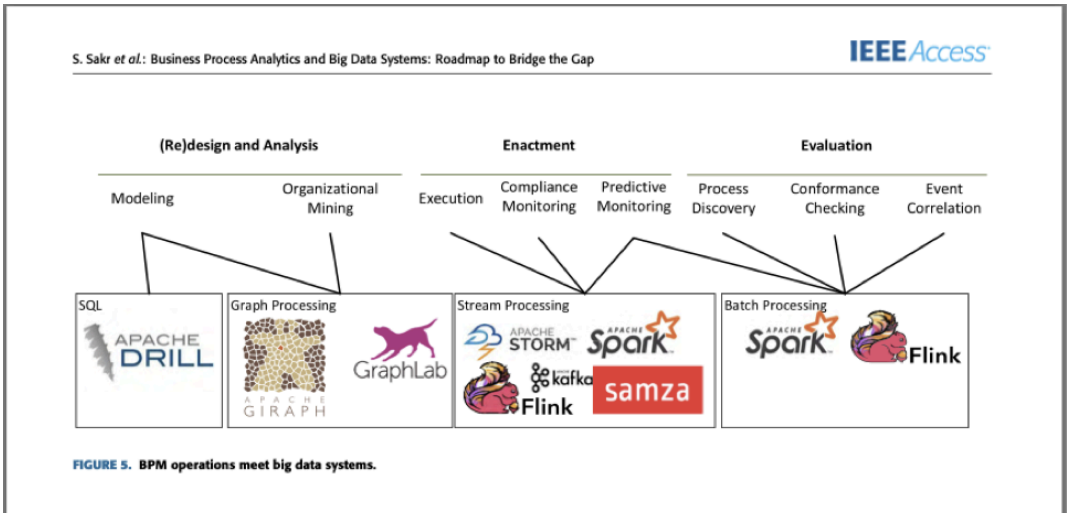

FIGURE 5. BPM operations meet big data systems.

To conclude, BPM can benefit significantly from leveraging big data systems and techniques in various stages of the BPM lifecycle. Data scales perceived the limitations of big data in the processing system and implantation of Business Process.

FILES 3 :

- Name: Pinto Frias Adriana

- Name of your Level 1: Eliott Laine

- Paper title: Business process analysis and optimization, beyond reengineering

- Source (e.g. scholars.google.com): scholars.google.com

- Keywords specific to the paper:

IDEF, NLP, XLM-based, Petri-Nets, Diagrammatic Models, Mathematical models, Business process languages

- Al model used (e.g. Neural network, etc.) : How do they contribute the idea proposed by the paper?

Neural network is used with a transformer model for natural language processing used for modeling language.

Also, NLP (natural language processing) are used with XML-based to analyze language, understand and automate the modeling languages.

- Supported by a software application? (If yes, provide more details)

IDEF : as a diagrammatic model for capturing and communicating knowledge but also observing and simulating business process analysis.

Petri-Nets : as a mathematical modeling language for description, simulation, performance analysis and so on.

- Summary of the main contributions (use text paragraphs, tables and if necessary, figures):

I. Introduction :

Reengineering is used for optimization and automated improvement using quantitative measures of performance.

II. Business process models, a novel classification :

Introduction of the most representative business process modeling techniques and classifies them in 3 groups based on characteristics.

- Diagrammatic models, simple and communicative :

These techniques use graphical representations. They are use-full for fast and informal process representation but there is a deficiency for complex products that led to the creation of graphical process modeling technique used to apply systems en engines software such a : IDEF. But they also know deficiency of formal semantic and for the visual inspection subjective.

- Formal/ Mathematical models, consistent but complex :

It's a process concept defined rigorously and precisely. Mathematics can analyze them but the problem is the lack of formal methods to support design processes.

For example : Petri net is a mathematical modeling language that uses visual representation, notation and mathematical representation.

So, building BPM analysis tools can be analyzed and extract quantitative information about the process.

- Business process languages, new and executable :

Business process languages are the most remote generation of BPM, used for process automation. XML-based is a Petri-Net model used for structured documents in a format. Business process modeling languages (BPML) products create interaction of the flows of a business process in an executable form.

III. Business process analysis, an overview :

Identification of the different types of business process analysis techniques and classify them in a similar way.

BUSINESS PROCESS MODEL	MODELLING SETS	BUSINESS PROCESS ANALYSIS TYPES
Idef	diagrammatic models	observational simulation
Rad's	diagrammatic models	observational performance analysis
Petri-nets	diagrammatic models mathematical models	observational validation vérification performance analysis simulation
Mathematical models	mathematical models	performance analysis simulation
Business process languages	business process languages	performance analysis simulation

IV. Business process optimization, beyond reengineering :

Justify the necessity to move from business process improvement to structured optimization and identify the sacred optimization approaches in the literature.

To represent business processes they also support the need for developing querying, linear programming, and simulation to select the optimal design.

- Business process languages : used for executable models with optimisation potential.
- Mathematical models : used for algorithmic objectives-oriented approaches and graph reduction.
- Diagrammatic models : used for graph reduction and unstructured modifications.

V. Overview and discussion :

To conclude, classifying business process models based on their mathematical, diagrammatic and language characteristics give us a wide selection of references regarding business process modeling, analysis, and optimization. With the help of process modeling techniques such as IDEF and Petri.

FILE 4 :

- Name: Pinto Frias Adriana

- Name of your Level 1: Eliott Laine

- Paper title: Deep learning in business analytics and operations research: Models, applications and managerial implications

- Source (e.g. scholars.google.com): scholars.google.com

- Keywords specific to the paper:

Neural network, Weight optimization, NP-hard, Weight decay, Batch normalization, Convolutional neural network, Recurrent neural network, Deep-embedded, Embeddings , Neurons, Grate recurrent unit, Long-short term memory, Multi-layer perceptron

- Al model used (e.g. Neural network, etc.) : How do they contribute the idea proposed by the paper?

Neural network : used to do a lot of tasks used with interconnected neurons

Deep neural network : used for NLP by analyzing raw data

Deep learning : it is a set of machine learnings used to analyze a large amount of data

Recurrent neural network : used to handle sequential data, mainly used for NLP.

Convolutional neural network : related to tasks for computer vision.

- Supported by a software application? (If yes, provide more details)

NP-hard : used to optimize neural network

- Summary of the main contributions (use text paragraphs, tables and if necessary, figures):

I. Introduction :

Deep neural networks allow the network to learn complex schemes and representations from the input data. Each layer connects nodes and performs mathematical operation of the input data. 1st layer → 2nd layer → output layer

As deep learning is for predictive analysis in business analytics, improvement is expected by deep learning for prediction performance.

II. Mathematical background, from neural networks to deep learning :

Predictive analysis : Exploiting features to predict unknown variables help to forecast future production needs and anticipate. For this we can use methods like : decision trees, support vector machines, neural networks or deep learning accompanied by optimization strategies.

Neural networks : Is a single-layer neural network that collects units, also called neurons. Artificial neurons receive input from other neurons and produce it into output.

Deep neural networks : Is a mathematical specification of multiple layers. It is difficult to choose an adequate number of layers and neurons.

Model estimation :

- Weight optimization : The optimization in deep learning is similar to the general setting in predictive analytics where the loss is minimized.

NP-hard are used to optimize a neural network, using gradient-based numerical optimization to produce the correct output.

The learning rate ν controls the step size during the optimization process

- Regularization : Challenge the number of parameters in the model. It elaborates complex relationships in the data and a higher number of free parameters.

The weight decay : penalize large weights, dropping out random numbers of neutrons.

The batch normalization : it's the normalization of the output.

Advanced architecture : overview of common networks architecture in deep learning.

ARCHITECTURE	TUNING PARAMETERS	DATA STRUCTURE

Dense neural network : - Multi-layer perceptron	Activation function, number of layers and neurons → simple replacement of traditional models	- Feature vectors of fixed length
Convolutional neural network	Number and width of convolution filters → image or speech recognition	High-dimensional data with local dependencies
Recurrent neural network - Long short-term memory (LSTM) - Gated recurrent unit	Number of layers and neurons → Text classification, time series forecasting, text mining	- Sequential data - Sequential data

- Convolutional neural network (CNN) : take advantage of a pyramid structure to first identify concepts before passing these concepts to the next layer
Recurrent neural network : used to handle sequential data. It is mainly used for NLP (natural language processing to make predictions and generate output.

III. Methods and materials :

Deep neural networks can handle in its raw form without the manual need.
For example : Deep-embedded designed to handle variables and translate it into a dense representation. This transformation facilitates stable optimization of the neural network even for a large number. Embeddings, also used to handle variables, are modeled by a simple neural network and can even be optimized and can learn more effectively. Also, the fact of capturing relationships between data can improve the performance on the primary task.
So, a higher optimization time for neural networks is compensated by a higher prediction performance.

IV. Numerical result :

A subset is used to evaluate how the number of training observations effects model performance. However, this experiment supports our statement that large datasets are needed to train deep neural networks because they encourage results.

V. Conclusion :

To conclude we can say that, deep learning methods improve prediction and performance for data-analytics models but they are complex architecture and require vaste parameters. Then, customized network architecture is recommended. Also, deep learning methods can create value for firms for organizations, business units, domains and decision-making.