UNIVERSITY OF TARTU
Faculty of Science and Technology
Institute of Computer Science
Computer Science Curriculum

Amey Chandrakant Darekar

Predicting Next Best Action(s) To Improve Sales Metrics For Pipedrive Customers

Master's Thesis (30 ECTS)

Supervisor(s): Mahmoud Shoush, Msc

Tartu 2023

Predicting Next Best Action(s) To Improve Sales Metrics For Pipedrive Customers

Abstract:
Predictive process monitoring (PPM) techniques exploit the full potential of historical event log data by applying data mining and machine learning methods to predict future process behavior, such as predicting or recommending the next best activity (or action). Modern techniques for recommending the next best action, particularly those using Deep Neural Networks (DNNs), have achieved near-perfect accuracy in predicting future process behavior in business environments. Despite this, since these techniques do not take into account Key Performance Indicators (KPIs), the metrics used by businesses to measure process performance making these techniques are limited in their ability to improve business processes in real-world applications. Process simulation has been used in the past to incorporate KPIs to optimize the process flow of business transaction activities, but this technique is limiting when there is a lack of definitive outcomes for action. In such cases, attempts to use process simulation alongside decision support for controlling action flows often yield unfavorable outcomes. We propose an approach inspired by business process optimization that relies on the probabilistic distribution of action sequences to predict the next best action(s). We attempt to implement this technique by taking into account KPIs that optimize the success rate of the sales transactions, using real-world event logs extracted from Pipedrive CRM. We also conducted experiments with heuristic search strategies to measure their usefulness when paired with our proposed strategy. We compare the performance of our proposed framework with the traditional control-flow simulation-based technique.

Keywords:
Prescriptive Business Process Management, Key Process Indicators, Customer Relationship Management, Process Mining, Deep Learning, Activity Recommendation

CERCS:
P176 Artificial intelligence

Järgmiste parimate meetmete ennustamine Pipedrive'i klientide müügimõõdikute parandamiseks

Lühikokkuvõte:
Ennustava protsessi jälgimise (PPM) tehnikad kasutavad ajalooliste sündmuste logiandmete kogu potentsiaali, rakendades andmete kaevandamise ja masinõppe meetodeid, et prognoosida protsessi käitumist tulevikus, näiteks ennustada või soovitada järgmist parimat tegevust (või tegevust). Kaasaegsed tehnikad järgmise parima tegevuse soovitamiseks, eriti need, mis kasutavad Deep Neural Networks'i (DNNs), on saavutanud peaaegu täiusliku täpsuse ärikeskkondade tulevase protsessikäitumise ennustamisel. Vaatamata sellele, kuna need tehnikad ei võta arvesse tulemuslikkuse põhinäitajaid (KPI), on

näitajad, mida ettevõtted kasutavad protsessi tulemuslikkuse mõõtmiseks, muutes need tehnikad piiratud nende võimega parandada äriprotsesse reaalsetes rakendustes. Protsessisimulatsiooni on varem kasutatud KPIde kaasamiseks, et optimeerida äritehingute protsessivoogu, kuid see meetod on piiratud, kui puuduvad lõplikud tegevuse tulemused. Sellistel juhtudel annavad katsed kasutada protsessi simulatsiooni koos otsuste toetamisega meetmete voogude kontrollimiseks sageli ebasoodsaid tulemusi. Pakume välja lähenemisviisi, mis on inspireeritud äriprotsesside optimeerimisest, mis põhineb tegevuse järjestuste tõenäolisel jaotusel, et ennustada järgmist parimat tegevust. Püüame seda tehnikat rakendada, võttes arvesse KPI-sid, mis optimeerivad müügitehingute edukust, kasutades Pipedrive CRM-ist saadud reaalmaailma sündmuste logisid. Samuti viisime läbi eksperimente heuristiliste otsingustrateegiatega, et mõõta nende kasulikkust, kui need on seotud meie pakutud strateegiaga. Me võrdleme meie pakutud raamistiku jõudlust traditsioonilise kontrollivoolu simulatsioonil põhineva tehnikaga.

Võtmesõnad:
Prescriptive Business Protsess Management, Peamised Protsessi Indikaatorid, Kliendisuhete juhtimine, Protsessi Kaevandamine, Süvaõpe, Tegevussoovitus

CERCS:
P176 Tehisintellekt

Contents

List of Figures

List of Tables

1 Introduction

In the dynamic world of modern business, the quest to optimize operations, enhance customer experiences, and drive growth has led organizations to embrace innovative technologies and strategies. By employing process mining techniques, businesses can uncover intricate patterns, trends, and correlations buried within their data. This enables the systems to better predict customer behavior, preferences, and needs, facilitating the delivery of beneficial insights that can drive the revenue to its full potential. Businesses increasingly rely on data-driven decision-making to enhance operations and overall performance.

Predictive process monitoring (PPM), which deploys machine learning algorithms to foresee future outcomes based on historical data, has emerged as a potent instrument for achieving such objectives [TVRD17]. Notably, the utilization of context-aware machine learning models has exhibited immense efficacy in this domain [AGSD22]. Deep neural network (DNN)-grounded predictive models have showcased remarkable accuracy in anticipating business process behavior [CDGR19]. While these results are impressive, these approaches often lack the influence of the metrics tied to business processes that steer the success-driving key performance indicators (KPIs), consequently falling short of delivering optimal outcomes for the business stakeholders and the end users. For instance, forecasting whether a specific task should be undertaken next to finalize a business transaction lacks substantial utility in terms of overall performance unless it's fine-tuned for relevant KPIs like cost reduction or enhanced customer satisfaction.

On the other hand, prescriptive techniques including the Prescriptive business process monitoring (PrPBM) technique to recommending next actions proposed by Weinzierl et. al. [WDZM20], advocate process simulation as a successful strategy to integrate KPIs into predictive business process models. While these techniques prove effective in manufacturing, supply chain management, and customer support/service processes, their application within dynamic process flows is hampered by inefficiencies arising from variability, uncertainty, and indeterminate outcomes.

1.1 CRM and Pipedrive

Pipedrive is a customer relationship management (CRM) software that is designed to help businesses of all sizes manage their sales pipeline and streamline their sales processes. It offers a range of features and tools, including lead capture and management, deal tracking and management, pipeline visualization, and sales reporting and analytics. The application provides a user-friendly interface and customizable features making it easy for businesses to track and manage their sales activities and improve their overall sales performance.

While many different features and extensions can be added to extend the experience, with different views the core of the application are pipeline and Details view for deals.

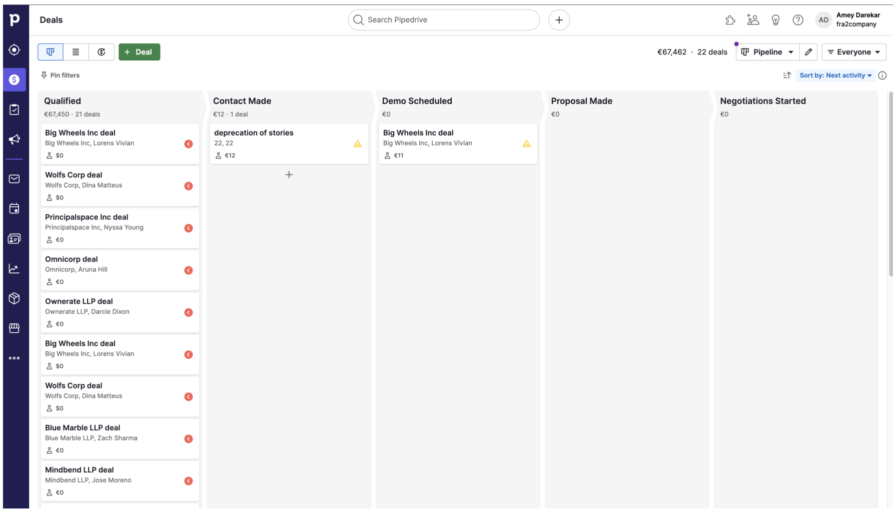

Figure 1. Deal pipeline in Pipedrive

The deal pipeline view, provides an overview of the different stages a deal can be in. A stage is a descriptive status of a deal with the intent of organizing deals into logical subgroups. These stages are fully customizable by the user, depending on their personal business process flows, but a deal is intended to resolve into one of two states - won or lost deals. In the deal Details view, user can add or update information about the deal, and schedule and plan actions or activities such as calls, email and meetings for a specific deal.

1.2 Motivation

Recommendation systems are tailored to serving specific agenda aligning with the user's needs and objectives. Netflix, Spotify and Youtube all suggest their hosted content to the users based on their likes, dislikes and interests. This focuses on user engagement as one of the important performance metric. Metrics such as these allow validating if the solution provides value to the user. KPI tracking is essential in Sales CRMs (Customer Relationship Management) because it provides businesses with a way to measure their performance against specific goals and objectives. KPIs are quantitative metrics that help businesses evaluate their progress towards achieving their sales targets and identify areas for improvement.

In the context of a recommendation systems for business processes, an activity refers to a specific step or action that is recommended to be taken as part of a particular workflow or process. For example, in a sales process, an activity may include sending a follow-up email to a customer or scheduling a call with a potential lead. The recommendation system would analyze relevant data, such as past interactions with the cus-

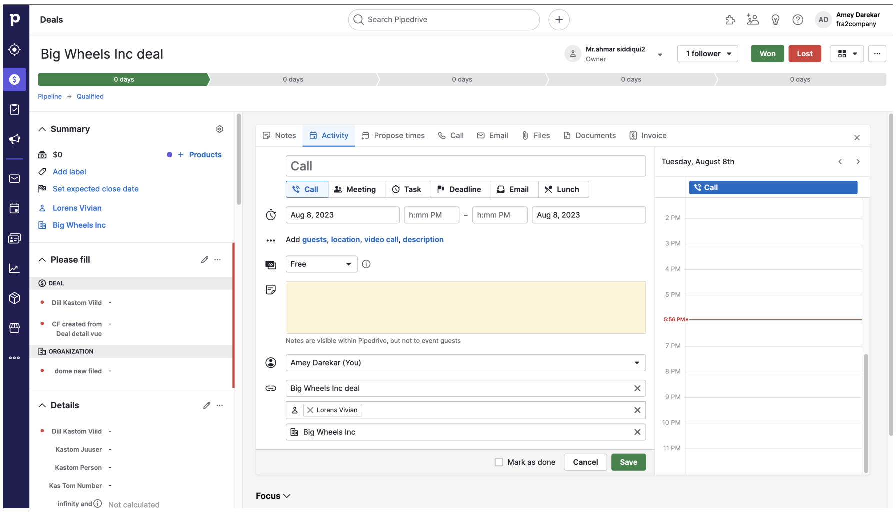

Figure 2. Schedule activity in deal detail view

tomer or lead, and suggest the most effective next activity to take based on that data. By providing personalized recommendations for specific activities, businesses can optimize their processes and improve their KPIs. Building an activity recommendation system can help mitigate several problems in terms of improving the KPIs. One common problem observed and reported across Pipedrive users is lack of personalization in their sales processes, especially when working with large volume of simultaneous open deals, which can lead to lower conversion rates and customer engagement. Another challenge is the difficulty in identifying the most effective next actions to take, particularly in complex sales processes. Additionally, businesses may struggle to optimize their sales processes and allocate resources effectively. By building an action recommendation system, businesses can address these issues by providing personalized recommendations to the users, suggesting the most effective next actions to take, and optimizing sales processes to improve KPIs.

By providing targeted and personalized recommendations to sales representatives, an activity recommendation system can help them identify the most effective next steps to take with each deal or customer. This can allow maximizing efficiency of the sales team and improve the KPIs such as increased conversion rates, faster completion times and growth in revenue.

1.3 Contribution

Context aware approach has proven very useful for building recommendation models in business process. The aim of this thesis is to create a well fitting machine learning model that can predict the next best action to be performed on case by case basis for the deals in order to achieve better KPI metrics. For the initial scope, we aimed at improving the North star metric for sales activities, which is deal conversion rate. We attempt to improve probability of improving deal win ratio by providing activity recommendations. Steps to building this model includes event log extraction, preparation of activity traces, building and training of the recommendation model, and finally applying the model to predict new user events.

1.4 Thesis outline

This thesis consists of 6 chapters. Chapter 2 gives overview of related work and advancements in the field of PPM, and involvement of KPIs in the models proposed. Chapter 3 goes in depth on data extraction and preparation for training the recommender model. Chapter 4 is dedicated to elaborating our proposed approach and implementation of the predictive machine learning model. Chapter 5 expands on the experiments performed and results of the performance. Chapter 6 concludes the thesis, gives a summary on what was achieved, expected impact of incorporating KPIs in the recommendation system on sales representative behavior and productivity with our approach and also supplies some notes on potential future work with this topic.

2 Related work

This chapter gives an overview of related works proposed and in and around use of mining and predictive use cases in business processes and also presents the current approaches to predict activity recommendations in the context of business processes.

2.1 Preliminaries

Following part of the thesis introduces the commonly used terminologies and their meanings, and provides brief explanation of these concepts in the context of recommendation task in business process. These terms and their interpretations have been defined and discussed with some variations in many prior works, but we use adapted definitions from works of Polato et. al. about remaining time prediction [PSBdL14].

2.1.1 Item

The term Item denotes what is being recommended to the user. Normally item is a specific type of entity or action being recommended, (e.g. a movie or an artist). When representing an event log, an item is a well-defined step in some workflow/process. In context of Pipedrive's Activities entity these are certain actions performed by the user with or without participation of the client that attempts to progress the deal towards an outcome. This can be a call, an email or an in person meeting at the client's workplace. Users can create custom activities designed for the scope of their company's usage that can has more meaning for them, such as "send estimate costs to the client via email", which extends the context of default email type of activity. For the simplicity of understanding, we elected to use only default types of activities which are common for all customers of Pipedrive. While input represents an individual activity performed by the users, the outcome recommendation is the type of action to be used when scheduling next activity.

2.1.2 Timestamp

Timestamp is the date and time documented at which something worthy of note happened. System can provide one or more relevant timestamps associated with each event, such as timestamp for when meeting scheduled, when meeting started and when meeting was concluded.

2.1.3 Resources

This includes any and all details that enrich the information about the event such as involvement of any business assets in the execution of an activity. This can be represented either as a single key-value collection, or in separate columns depending on the data cleaning and processing operations applied.

2.1.4 Event

An event is an individual, smallest unit of information in an event log. At the minimum, it should contain the activity and timestamp. Optionally, it may have information associated resources, life cycle, and other data.

2.1.5 Case

A related set of events denoted, and connected, by a unique identifier where the events can be ordered. A transaction between user and the client being tracked with the business process flow in this context is considered as a unique case.

2.1.6 Event Trace

Event trace is a distinct selection of case activities within an event log where each event ties to single process sequence, and events are ordered in the order of occurrences.

2.2 Predictive Business Process Models

2.2.1 Prediction of case outcome

One of the most important and commonly measured metrics of sales process is the number of successful or won cases. The success ratio of the transactions gives the good insight to the efficiency of the business process and the performance of the sales team. Most sales teams set their primary KPI goal for to "Win more cases". A machine learning model predicting win probability of a case can assist the user to take appropriate actions to manage the resources to focus on improving the undesirable outcomes. The insight into likeliness of an outcome being favorable or not can allow the users to decide on prioritising and execute favorable transactions, or provide more resources towards transactions predicted to have a bad outcome depending on the significance of the same. In the scope of this thesis, we do not address the problem of outcome prediction directly, but propose a recommendation architecture to improve the probability of the case outcome and in effect, improve the case win ratio.

2.2.2 Prediction of completion time

The KPI "Win cases faster" refers to a key performance indicator used to measure the efficiency and effectiveness of a sales representative or team in successfully closing deals or sales opportunities in a shorter time frame. It assesses the ability of the sales process to move swiftly from initial engagement with a potential customer to the final conversion and closure of the sale. The quicker the sales team can close deals, the more efficient and productive their sales process is perceived. A predictive model that predicts completion time of the deals can empower sales teams to work more efficiently, make strategic decisions, and prioritize their efforts to close deals faster. This, in turn, results to increased revenue, enhanced customer satisfaction, and a competitive advantage in the market.

2.3 Recommendation Systems

Recommendation systems are information filtering systems providing a personalized item recommendation to a user in a service environment that can hold or collect various data. Information filtering, which is mainly used in recommendation systems, is tailored to the user's preferences or suggested only items judged to be useful to the user [KLPC22]. "Item" is the general term used to denote what the system recommends to users. An RS normally focuses on a specific type of item (e.g., movies or news articles) [RRS22]. In order to provide these recommendations, it is necessary to gather user's implicit and explicit data for item recommendation. This data should be utilized to mine for insights and usage patterns and suggest meaningful actions. Furthermore, the process should be refined based on newly generated data, feedback on the recommendations, and correlation evaluation between users, etc as a periodic process.

Depending on the use case, there are two major types of recommendation systems. In recommendation systems dedicated towards personalizing the recommendations for consumers where objective is suggesting novelty or exploration, such as in music in applications such as Spotify, shows and movie recommendations in Netflix, or suggesting posts, groups and connections in social media platforms. For these type of recommendations, the model is formulated as matrix completion problem, where each item is recommended to the user only once. On the other hand are cases where a certain item can be recommended multiple times, as the application of these suggestions can be carried out multiple times. For such models, the aim of is to look for behavioral patterns, to optimize the effectiveness of the actions, and improve engagement.

2.4 Sequence aware recommendation systems

Tasks approached with Sequence aware recommendation systems are quite different than matrix completion problems in many ways. The key difference between these types of models is that ordering of objects becomes relevant for both inputs and outputs. The inputs are represented in the forms of interaction or event logs and produced output is either single prediction or an ordered list of actions.

Massimo et. al. [QCJ18] describe categorisation of sequence-aware recommendation tasks to achieve 4 types of goals, according to application scenarios:

1. Context Adaptation
2. Trend Detection
3. Repeated Recommendation
4. Consideration of Order constraints and Sequential patterns

2.4.1 Context Adaptation

In many domains, the reliance of recommended item on user's preference alone is not enough, but requires insight into process context, short-term intents and interests. Context- aware recommendation systems take such additional types of information into account. Typical contextual factors often include the user's geographical position, the current social or environmental context if applicable, or even temporal context such as the time of day. This representation of the context is a predefined set of observable variables.

2.4.2 Trend Detection

The detection of trends is yet another possible insight a sequence-aware recommendation systems can provide to the users. A sequential log of events or item interactions can allow extraction of trends on global and individual scale for benefit of recommendation process. Global trends can help understand popularity patterns and devise long term strategies, where individual trends can allow identification of influences and personalising the recommendation aligned to natural interests.

2.4.3 Repeated recommendations

In application domains, recommending items previously known or applied again, can prove useful. This type of scenario is not considered in traditional matrix completion problem. This approach is useful in two distinguish cases while formulating the problem. One is to recommend action based on repeated user behaviour patterns, with the intention of enhancing user experience. This was explored as a use case for recommending application to be used in smart phones in [NSD13]. Second case for using repeated recommendations is as a method as a reminder to interests or actions as a reminder. This could be items user have used or shown interest to in the past. This is a common scenario for e-commerce businesses and social media platforms, where recommendation is recently visited items can be revisited.

2.5 Advantages of Context in recommendation task

The objective of content based recommendation tasks is to capture user's general or long-term preferences. The items recommended usually are static preferences, and remain relevant for long intervals. However, in case of fast evolving state of the transaction, such approach falls short. In the Context aware recommendation task approach, incorporating contextual information such as details about the current transaction, participating users, previously carried out actions, involved clients, time since previous action, time of the day for new action prediction are to be considered in the prediction flow. This allows the recommendation model to provide more personalized and relevant recommendations to the user. For context aware recommendations in business processes, following advantages can be observed:

- Improved personalization: Context-aware systems can provide more personalized recommendations by taking into account the user's current situation and preferences in that context.
- Broader range of recommendation: Context-aware systems can recommend a wider range of items that are relevant to the current context of the process, rather than simply recommending items that are similar to what the user has liked in the past.
- Enhanced user engagement: By providing more frequent, personalized and diverse recommendations, context-aware systems can increase user engagement with the system, leading to increased satisfaction and loyalty.
- Accurate recommendations: By including contextual information, context-aware systems can provide more accurate recommendations that reflect the user's current needs and preferences.

2.5.1 Long short-term memory neural networks

Use of deep neural network techniques has been practicable in the context aware recommenders, as cited in [AGSD22], [QCJ18], and [WZB+20]. This approach allows efficient mining of the sequential patterns and handle temporal dependencies of the sequence prediction thus improving overall efficacy. Most commonly used model amongst these techniques the of Recurrent Neural networks (RNNs) [LBH15] is Long Short-Term Memory (LSTM) network [HS97], where short-term and long-term rewards to ensure that the recommended activity is not only optimal for the current context but also contributes to achieving the user's long-term goals.

2.5.2 Use of Reinforcement learning frameworks

Reinforcement learning (RL) has been used in recommendation systems to make personalized recommendations to users by learning from their feedback. The role of RL in recommendation systems is to learn the optimal policy for recommending items to users based on their past behavior and the context in which they are interacting with the system. In RL-based recommendation systems, the system observes the user's actions and rewards and uses them to update its recommendation policy, with the goal of maximizing the user's satisfaction or engagement with the system. This is in contrast to traditional recommendation systems, which use supervised or unsupervised learning techniques to make recommendations based on patterns in user behavior or item features.

The benefits of using RL in recommendation systems include the ability to learn from user feedback in real-time, the ability to make personalized recommendations based on individual preferences and context, and the ability to adapt to changes in user behavior over time. RL-based recommendation systems also have the potential to discover new items that users might not have otherwise found, leading to increased user satisfaction and engagement. However, RL-based recommendation systems also have some limitations, including the need for large amounts of data to train the system, the challenge of balancing exploration and exploitation, and the potential for the system to become biased towards certain items or types of users if the training data is not diverse enough.

Overall, RL has shown promise as a powerful tool for building more effective and personalized recommendation systems, but it is important to carefully consider the benefits and limitations before implementing an RL-based approach in a specific application.

In [AGSD22] Agarwal et. al. propose Goal-Oriented Next Best Activity Recommendation using RL based framework to provide a personalized next best activity recommendation aligned to the user's goals. The proposed approach utilizes Deep Neural Network (DNNs) with combined use of Generative adversarial network (GAN) and Long Short-term memory (LSTM) techniques. Based on the evaluation of real-world dataset of insurance claim adjusters' activities, authors compare the performance of their RL algorithm with two baseline methods: random recommendation and rule-based recommendation. The results show that the proposed solution outperforms the baseline methods in terms of the number of goals achieved and the efficiency of the recommended activities. The paper concludes that the proposed RL-based approach can effectively learn a personalized NBA recommendation policy that is aligned with the user's goals.

2.6 Supporting predictions to improve goals and metrics

Use of state-of-the-art deep neural networks (DNNs) perform exceptionally well in recommending actions accurately for the data, it does not help ensuring the process performance is affected positively in the business context. Therefore, the resulting next most likely activity predictions can be less beneficial in practice. Prescriptive business process monitoring approach to next best action recommendations (PrBPM), proposed by Weinzier et. al. [WDZM20] assesses predictions regarding their impact on the process performance (typically measured by KPIs) to prevent undesired process activities by recommending actions. This approach implements business process simulation with use of DCR graphs to ensure control-flow conformance of recommended actions. The authors claim that the proposed approach outperforms the KPI (lead time) fulfillment baseline.

3 Data Collection and preprocessing

As stated in Motivation, the objective of thesis is to provide recommendations of Pipedrive Activities to improve sales KPIs for Pipedrive users.

3.1 Data Overview in Pipedrive

Pipedrive's customers use following types of objects to store, access and interact with their business process actions.[1]

- **Person** - Person entity in Pipedrive application represents individual contact representative of a client of Pipedrive user. User can enter client's contact details, important information and notes about the client by creating or editing respective Person entity. Every persons can be have multiple Deals and Activities associated with them.

- **Organization** - Organization represents the company or entity their clients (Persons) belong to. An organization can have multiple Deals and People related to them. A Person cannot belong to more than one Organizations.

- **Deals** - Deals represent transactions Pipedrive user is conducting with their clients. Deals can be linked with clients at Organization or Person level or both. Every deal belongs to certain stage in the Pipeline, which represent the state of the deal at a given time. Every created deal is considered ongoing/open and can be closed by updating the status to either won or lost.

- **Pipeline** - Pipeline is the flow of business transaction flow broken down into stages as per domain requirement. Stages of pipeline are customizable and can be added or removed from the pipeline based on user's needs. Deals can move forward or backwards in the pipeline by updating stage of the deal.

- **Activities** - Activities represent the actions user can schedule and conduct towards closing a sale. It can be a phone call, a lunch meeting, an email or any other event to be scheduled with a contact. You can schedule activities in relation to a person, organization or a deal.

3.2 Data Processing and Warehousing

With over 100,000 companies (and even more users) using Pipedrive for managing their sales activities, a large amount of data is being generated inside Pipedrive on a daily

[1] https://support.Pipedrive.com/en/article/how-is-Pipedrive-data-organized

basis. This data comes from various sources such as user actions, usage metrics tracing, marketing campaigns and support activities. This data can be used to serve many possible purposes depending on meaning derived from different stakeholders.

To make process of processing and using this data easy, robust and streamlined, Pipedrive has developed its very own Data Stack, that caters all sort of stakeholders from a universal source. Following sections provide overview of the Data stack and role for each of its major components.

3.2.1 Data Warehouse

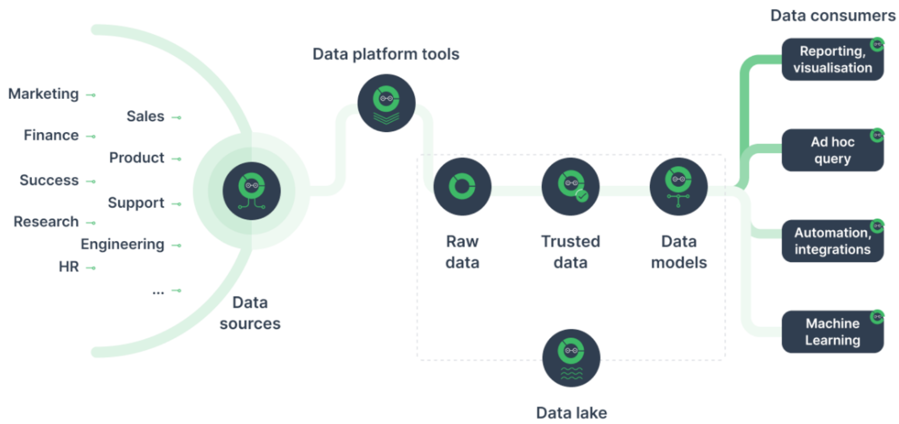

Figure 3. Data overview in Pipedrive

Data Warehouse (DW) in Pipedrive is an umbrella of tools and methodologies implemented and maintained by dedicated team of Engineers and analysts. The Data team is responsible for the processes of storing generated data, ensuring data available is processed properly to achieve granularity, and in anonymized form where necessary. The team also controls who inside Pipedrive can access the data under the DW. Key components of DW are as following.

- **S3 Storage** - Amazon's S3 object storage service is used as staging temporary data, a primary storage for processed data source for preprocess pipelines and destination for outcome of Apache Spark jobs. Access to the Amazon provided services are protected as a need based access management, and protected by LDAP authentication.

- **Apache Spark** - Apache Spark is agile, in-memory data processing tool that allows stream or batch processing massive amounts of data efficiently. It is based on Hadoop's MapReduce programming model, consisting of two procedures Map (filter and sort data) and Reduce procedure which is a summary operation. Additionally, Spark also supports data analytics, SQL queries, Machine Learning and graph algorithms. [ZCF+10]

- **Apache Zeppelin** - Apache Zeppelin is an open source web-based computing platform, provides an integrated environment for data ingestion, data discovery, data analytics and data visualization and collaboration with an extended framework which allows different programming languages and data processing back ends to be plugged in. Currently it allows creation of collaborative documents with SQL, Scala, Python and more. It has the notebook and visualization capabilities to make big data more approachable and easier to understand. The Zeppelin's multi-purpose Notebook provides data ingestion, data discovery, data analytics and data visualization and collaboration. Its interpreter concept allows any language/data processing back end to be plugged into Zeppelin. Currently Apache Zeppelin supports many interpreters such as Apache Spark.

3.3 Data Schemas

Data in the S3 storage is split into 3 layers, based on data granularity and processing applied:

- **Datalake Bronze Layer** - This is data stored in the rawest form, and is used as raw input for Silver and Gold layers. The relevant schema names are prefixed by the source of the data.
- **Datalake Silver Layer** - In this layer, data from bronze layer is processed with automated cleaning and enrichment data flows. This layer also contains the primary Data lake schema dw. The tables in this schema can be broken divided into two type of layers:
  - Fact tables contain the quantitative information of each entity. These are used to have easy access to statistical information about the data itself, like daily active users in a company.
  - Dimension tables contain different attributes of business processes.
- DW Gold Layer - This layer contains cleaned and processed data at its most improved version and is most frequently used to perform analytical operations and draw statistical measures.

Having defined the layers from which data in Pipedrive can be retrieved, it is now possible to elaborate schemas of tables in DW used to query for extraction and preparation of data required for training the model. A detailed overview is added to the Appendix I. The data extracted from the model is as follows. As model's intention in initial stages is to generalise the recommendation process across users of Pipedrive from many companies clustered together based on region or operational domain,

- **company** - Information about the user company's industry domain and country, to extract more closely connected data.
- **deal** - This table contains information about deals belonging to the companies, along with IDs for various relations such as users, persons, organizations. The important features extracted from this tables are the deal's status representing if case represents successful or failed transaction and the creation and conclusion timestamps for the deal.
- **activity** - Information about individual activities carried out by the user, participants involved, timestamps for creation, start, estimated due date and actual completion timestamps.

3.4 Data selection and feature engineering

Chapter 3.3 elaborated on the tables where suitable event logs are to be queried. This section further elaborates on the conditions and filters applied in the data preprocessing in order to train and evaluate the machine learning model.

1. We filtered and selected deal cases with most similar time duration and number of activities to train a model that can fit best on about 90^{th} percentile of the cases. The deals that were part of pipeline with 5 to 8 stages were considered as ideal cases for the model training.
2. Only event logs with cases created and concluded in the year 2022 are considered. The recent time-frame ensures that the model will learn user behaviour patterns relevant to the relatively latest event data. Only cases which are marked as completed, with status as 'won' or 'lost' are considered for training the model, as we rely for this status to score prediction usefulness as a measure.
3. Only 10 activity types are considered which are available to all companies by default, We filtered out the deals that involved user created and/or customized activities from the data, as these types of activities were less frequently used, did not provide any additional insight or predictive support and suggesting one such action type specific to users in a company cannot be suggested to another user who does not implement same actions in their process flows.

4. Only, cases with at least 3 items in the event trace are considered, as smaller traces suffer the cold-start problem, where smaller event traces do not provide enough contextual information for case, and thus affects prediction scores adversely.

5. The included traces only contain successfully carried out actions, and contain only latest information in cases. All in progress and abandoned cases were removed.

3.5 Preprocessing and feature engineering

Chapters 3.3 and 3.4 have established the data available towards building the recommendation framework. This chapter describes the features engineered from the existing data, analyze the data with exploration techniques and preparation of the data as input to the machine learning model.

The most important step of the preprocessing stage was mapping the event names to a numeric value. Secondly, the events in the trace should be arranged by sequence of occurrence. Using timestamp of the events, it is possible to identify events sequence of completion. Additionally, for data analysis purposes, following features were also engineered:

1. Minutes, hours and days taken for event completion.

2. Total count of events in a trace.

3. Total number of days a case was active.

4. The time difference between completion of previous event and start of the current event.

3.6 Exploratory data analysis

Followed by the feature engineering, we attempted preliminary data exploration, to draw some insight for model construction. We performed seasonality check and clustering on the cleaned and feature engineered dataset, which showed certain criterion that suggested the possibility of loss in generalization and data leakage in the training stages. The extended event log was then processed with use of Action Rule discovery framework. This framework, as described in [RD06], provides action rules, set of actionable updates to features of the dataset, , to discover the influential features towards success of outcome of each case. The case context, such as user, company, case transaction value, are treated as action premise, and event features, along with the features engineered in previous case are treated as the candidates for suggesting actionable changes. For example, the framework may suggest a rule like "For user X, with transaction value 2000, and 2 meeting events, the probability of the case success increases when one or more 'call' type of events are performed after 3 days. This rule has confidence of 20% and support at 15%". The insight into action recommendations from this process is summarized below.

1. The factual information of the case shows very little if no correlation to the case outcome.

2. The increase in frequencies of specific event types (activities) provide the most significant influence to decision.

4 Technical Implementation

This chapter covers how data is used, with the use of proposed architecture, to implement sequence aware recommendation system, process of suffix prediction and evaluation strategy that incorporates the KPI information.

4.1 Training activity prediction model

Inspired from the PrBPM model, the architecture of our proposed system is split into two major components. First is the activity prediction model, a 'vanilla' DNN classifier with LSTM that recommends next activity for the current event trace. In second component, the recommended action is optimised to improve KPI, and predict the probability of success. A visualisation of the implemented approach can be seen in Figure 4.

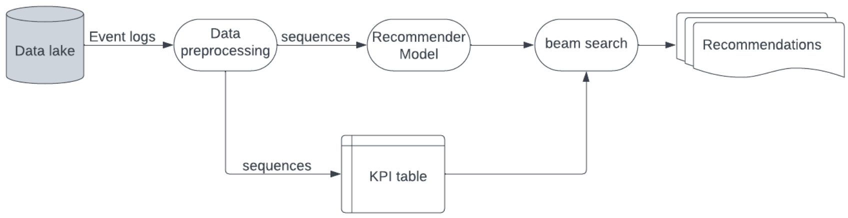

Figure 4. Architecture of the proposed activity recommendation system

The data extracted in Chapter 3.4 is used to train the activity prediction model. This data in the form of event logs is broken into smaller event traces along with the timestamps of the events. These traces are then split into prefixes and singular predicted suffix event. The model takes event trace e_k where k indicates most recent event in the trace e, and predicts the action of the next event e_{k+1}. The proposed model does not predict the start or end timestamps for the predicted event.

To train the model, the events are required to be encoded to model compatible structure. The varying length sequences are padded to an equal length, the event names, represented as categorical data is converted to one-hot representation and temporal data is converted to numerical form. The model architecture comprises two LSTM layers with dropout and batch normalization for stability. The input data consists of activity sequences represented as a 3D tensor. The model is trained using categorical cross-entropy loss and the Adam optimizer.

The code of the created solution is uploaded to a Github repository[2]. For each dataset in Table 1, the experimented were conducted on Jupyter[3] notebook instances on a AWS

[2]https://github.com/a3darekar/msc_thesis_best_next_action.git
[3]https://jupyter.org

SageMaker[4] with the following specifications:

- instance type: ml.t3.2xlarge
- vCPU: 8
- Memory: 32 GiB

The model was set to train for 40 epochs for each dataset, with an additional end condition being that early stopping is initiated when the accuracy on the validation set has not changed in the last 10 epochs.

4.2 suffix prediction

After activity prediction model is training is concluded, It is possible to implement the suffix prediction. First we prepare probability distribution for activity sequences with respect to case outcome. This distribution will be used to optimize prediction of the next action that will most likely aid in successful result of the case. Further optimization of the search space is carried out by using beam search [SVL14].

In this prediction context, The beam search algorithm employs the trained LSTM model to predict activity types in the iterative manner, exploring multiple candidate sequences and selecting the most promising ones based on scoring criteria. This approach facilitates the generation of diverse and contextually relevant sequences. The algorithm looks at n most probable predicted activities at any given time step, where n is a predetermined beam width. Then, the activities that are not in the top n candidates are discarded and the prediction for the next time step can begin. The process of prediction along a selected activity path is continued until the event that indicates the end of a sequence is generated or that event is not considered in the top n candidates anymore. When no partial predictions are left to be made, beam search algorithm finishes.

[4] https://aws.amazon.com/sagemaker

5 Results and analysis

5.1 Experimental setup

From the data criteria defined in Chapter 3, we extracted 2 datasets from Pipedrive's DW that fit the conditions for training the model. First dataset contains data from 5 companies operating in the same domain, and operate in the same geographical region. Second dataset is a subset of the first dataset, containing event logs generated by only one of the companies. The objective for creating this subset is to observe if collective model training benefits the performance for all users or individual company level model provides better predictions. During this process we took the precaution limiting the teams operating in similar sales process environment, with similar number of stages, participating sales representatives and clients on average. We cleaned and sanitized the data, processed engineered features as described in chapter 3. The datasets had around 41% and 37% positive case data respectively.

Dataset	**Number of event traces**	**Avg length of event traces**	**Max length of traces**
Collective dataset of 5 companies	13712	8	27
Single company dataset	10356	8.6	27

Table 1. Statistics of dataset used in experiments

Along with data extracted from Pipedrive's DW, we also used an external dataset[5] of loan application process from Business Process Intelligence (BPI) challenge from year 2017. We used this data to determine if the approach is useful to data that can follow process-flow patterns along with uncertain data like Pipedrive. The data consists of logs from the year 2016 pertaining 31509 applications from 149 users consisting 3 types of events describing various states in life-cycle of the application. For each uniquely identifiable event, the employee who caused the event is recorded, as well as a timestamp and life-cycle information. The data is provided in the form of the standard XES life-cycle events.

[5]https://www.win.tue.nl/bpi/2017/challenge.html

5.2 Evaluation metrics

We opted for two-fold evaluation metrics for the assessment of the model performance. The first metric was the "improvement over original trace probability", where we measured the improvements in success probability for each of the final predicted sequence in comparison with the original trace for each model. This metric gauges how much better our proposed solution performed compared to the original probability estimation for each trace in the event log data. By calculating the difference between the new predicted probability and the original probability, we could quantify the effectiveness of our model in enhancing the accuracy of outcome predictions. The second metric was the number of cases where the success probability for the trace sequence increased above a predetermined threshold. Here, we set a predefined threshold to determine the acceptability of the predicted probabilities. By counting the number of cases where the new probability surpassed this threshold, we gained insights into the model's ability to recommend actions that met or exceeded a specified level of certainty. This metric is particularly valuable in evaluating the model's performance in critical decision-making scenarios, where a certain level of confidence in the predicted outcome is crucial. The time taken for performing an action was taken into account for reporting the dynamics of sequence and success probability of the sequence but was not intended as focus of improvement in the proposed model structure.

Subsequent to conducting the initial experiments, we also focused on evaluating the model's performance using event log data from BPI 2017[vD17]. This loan applications process log, has a significant difference from the sales action logs in terms of trace structure and characteristics. The most glaring difference of the model is that each action performed in the process of a loan application updates the status of the application to a new state. Despite these differences, we were optimistic that the proposed solution would still yield favorable results and not suffer from any deterioration in probability outcomes. By incorporating the BPI 2017 event log data into our evaluation, we aimed to thoroughly assess the ability of our approach to adapt a state focused configuration and generalize to a broader array of applications.

5.3 Results and analysis

For each of the datasets, we conducted two sets of experiments with varying beam sizes of width 1, 2, 3, and 5. The first experiment implemented with only the beam search in process simulation phase, while second experiment utilized the lookup table to generate a prediction of success probability after each action recommendation. We set 5 as the beam width limit as initial tests did not yield better results and computational requirement for wider beam size results in increased processing time and memory overhead altogether. As the beam size increases, the likelihood of having already discovered the optimal option increases too. Consequently, the chances of additional improvement has diminishing returns.

The experiment results demonstrate a clear trend in success probabilities during process simulation. here's a gradual increase in success probability for cases with each increase in beam width, but this increase becomes significantly more pronounced when utilizing the KPI probability table (tables 2 and 3).

In both collective and single dataset scenarios, the incorporation of KPI-driven probabilities becomes evident. Success probabilities experience a noticeable surge, highlighting the efficacy of integrating KPI insights into predictions. This supports our goal of using KPIs to enhance decision-making in business processes. The difference in success probability progression underscores the advantage of our approach. This emphasizes the role of KPIs in improving predictive models for our specific context.

KPI Table used	**Average improvement to success probability with beam search width**			
	1	2	3	5
No	7%	11.68%	15%	19.85%
Yes	15.18%	27.82%	36.42%	37.20%

Table 2. Average improvement of success probability observed for Collective dataset of 5 companies

KPI Table used	**Average improvement to success probability with beam search width**			
	1	2	3	5
No	8.36%	13.92%	14.06%	21.19%
Yes	23.03%	30.83%	35.64%	38.30%

Table 3. Average improvement of success probability observed for dataset of single company

In the analysis of evaluation metric 2, we further confirm success of our technique as observed in tables 4 and 5. We see similar trends that for the experiments with the use of beam search show minor gradual improvements, and use of the KPI probability table provides much pronounced increase in the outcomes. This improvement is most likely due to beam search alleviating the greediness of Best first search by comparing more possible options. This improvement is notable as it stands, but in terms of practical usage it does not provide benefits significant enough for the added computational overhead it requires. In case of the second experiment, even with greedy approach, we see a significant increase in predicted improvement when using the KPI table for determining the next activity. As we increase the search space with increase in beam width, we observed a more significant increase in predicted improvements as well. In this case as well, we saw on average 40% increase in the predicted improvement.

KPI Table used	**Case improvement predicted with beam search width**			
	1	2	3	5
No	10%	12.30%	15%	19.85%
Yes	35.18%	37.98%	43.79%	48.86%

Table 4. Percentage of cases with outcome probability improved for Collective dataset of 5 companies

KPI Table used	**Case improvement predicted with beam search width**			
	1	2	3	5
No	9.19%	12.97%	14.97%	17.78%
Yes	35.09%	43.79%	46.77%	47.28%

Table 5. Percentage of cases with outcome probability improved for dataset of single company

Overall, our investigation with the BPI 2017 event log data provided valuable insights into the model's capacity to predict actions for cases when actions yield a certain outcome and state of a case changes accordingly these outcome scenarios. The recommender model worked with near perfect accuracy aligned with KPI probabilities in simple simulation experiments, and the involvement of KPI lookup table or beam search strategy did not provide any boost to the performance in this case.

6 Conclusions

In this study, we conducted an in-depth exploration of various recommendation model techniques and implementations, critically evaluating their strengths and weaknesses. Through analysis, we found that a significant gap exists in considering key performance indicators (KPIs) within the proposed techniques in the predictive business process tasks. Furthermore, we observed a prevalent reliance on past outcomes to inform subsequent actions, a methodology that does not seamlessly align with Pipedrive's sales activities log we wish to utilize for the task. This is attributed to the inherent ambiguity and non-quantifiability of activity outcomes in this context.

To address these limitations, we put forth a novel approach aimed at rectifying these challenges. This involves the computation of scores for suffix predictions, strategically aligned with the overarching objective of KPI enhancement. Inspired by the use of collaborative filtering techniques yet relatively novel in terms of sequential prediction domain or the context of business processes, we attempted to adapt the probabilistic matrix factorization approach in our technique for supporting suffix prediction. This adaptation serves the dual purpose of streamlining prediction computation while harnessing KPI-related insights to inform decision-making. By doing so, we mitigate the computational complexity associated with more intricate models, thus optimizing process simulation time without sacrificing the utilization of KPI-related information.

An extension to current model, a secondary objective to predict estimated time for action completion can be introduced to the problem space. In this scenario, the model could be trained to find an activity that helps achieving the second most important KPI for sales teams, 'win deals quicker'.

Looking ahead, our proposed method holds the potential for further expansion and refinement. One avenue for advancement involves the incorporation of a more robust case success probability methodology to replace the existing lookup probability table. The current KPI probability strategy only relies on the activity logs, and does not account other important factors involved in the process such as value of the case processed or number involved stakeholders. A dedicated prediction model designed to replace the current strategy would enhance the accuracy of predictions and bolster the reliability of the overall framework. Additionally, there lies an opportunity to reevaluate the benefits of the beam search strategy, potentially unlocking even more efficient and effective predictive capabilities within our approach. Secondly, the model should be tested and fine-tuned based on real world implementation, to improve further than the current process simulation method.

In conclusion, this study not only sheds light on the limitations of existing recommendation model techniques concerning KPI integration and uncertain outcomes but also introduces a novel and pragmatic solution to address these shortcomings. By embracing the principles of probabilistic matrix factorization and strategically aligning prediction scores with KPI improvement, our approach pioneers a more holistic and impactful method for optimizing business processes. As the realm of data-driven decision-making continues to evolve, our findings offer a valuable contribution to the ongoing discourse on enhancing the utility and applicability of recommendation models in real-world business contexts.

References

[AGSD22] Prerna Agarwal, Avani Gupta, Renuka Sindhgatta, and Sampath Dechu. Goal-oriented next best activity recommendation using reinforcement learning, 2022.

[CDGR19] Manuel Camargo, Marlon Dumas, and Oscar González-Rojas. Learning accurate lstm models of business processes. In Thomas Hildebrandt, Boudewijn F. van Dongen, Maximilian Röglinger, and Jan Mendling, editors, *Business Process Management*, pages 286–302, Cham, 2019. Springer International Publishing.

[HS97] Sepp Hochreiter and Jürgen Schmidhuber. Long Short-Term Memory. *Neural Computation*, 9(8):1735–1780, 11 1997.

[KLPC22] Hyeyoung Ko, Suyeon Lee, Yoonseo Park, and Anna Choi. A survey of recommendation systems: Recommendation models, techniques, and application fields. *Electronics*, 11(1), 2022.

[LBH15] Yann LeCun, Y. Bengio, and Geoffrey Hinton. Deep learning. *Nature*, 521:436–44, 05 2015.

[NSD13] Nagarajan Natarajan, Donghyuk Shin, and Inderjit S. Dhillon. Which app will you use next? collaborative filtering with interactional context. In *ACM Conference on Recommender Systems (RecSys)*, oct 2013.

[PSBdL14] Mirko Polato, Alessandro Sperduti, Andrea Burattin, and Massimiliano de Leoni. Data-aware remaining time prediction of business process instances. In *2014 International Joint Conference on Neural Networks (IJCNN)*, pages 816–823, July 2014.

[QCJ18] Massimo Quadrana, Paolo Cremonesi, and Dietmar Jannach. Sequence-aware recommender systems. *ACM Comput. Surv.*, 51(4), jul 2018.

[RD06] Zbigniew W. Raś and Agnieszka Dardzińska. Action rules discovery, a new simplified strategy. In Floriana Esposito, Zbigniew W. Raś, Donato Malerba, and Giovanni Semeraro, editors, *Foundations of Intelligent Systems*, pages 445–453, Berlin, Heidelberg, 2006. Springer Berlin Heidelberg.

[RRS22] Francesco Ricci, Lior Rokach, and Bracha Shapira. *Recommender Systems: Techniques, Applications, and Challenges*, pages 1–35. Springer US, New York, NY, 2022.

[SVL14] Ilya Sutskever, Oriol Vinyals, and Quoc V Le. Sequence to sequence learning with neural networks. In Z. Ghahramani, M. Welling, C. Cortes, N. Lawrence, and K.Q. Weinberger, editors, *Advances in Neural Information Processing Systems*, volume 27. Curran Associates, Inc., 2014.

[TVRD17] Niek Tax, Ilya Verenich, Marcello La Rosa, and Marlon Dumas. Predictive business process monitoring with LSTM neural networks. In *Advanced Information Systems Engineering*, pages 477–492. Springer International Publishing, 2017.

[vD17] Boudewijn van Dongen. Bpi challenge 2017, 2017.

[WDZM20] Sven Weinzierl, Sebastian Dunzer, Sandra Zilker, and Martin Matzner. Prescriptive business process monitoring for recommending next best actions. 09 2020.

[WZB+20] S. Weinzierl, S. Zilker, J. Brunk, K. Revoredo, A. Nguyen, M. Matzner, J. Becker, and B. Eskofier. An empirical comparison of deep-neural-network architectures for next activity prediction using context-enriched process event logs, 2020.

[ZCF+10] Matei Zaharia, Mosharaf Chowdhury, Michael J. Franklin, Scott Shenker, and Ion Stoica. Spark: Cluster computing with working sets. In *Proceedings of the 2nd USENIX Conference on Hot Topics in Cloud Computing*, HotCloud'10, page 10, USA, 2010. USENIX Association.

Appendix

I. Table Schemas

I.1 company

Column	Data type	Description
id	int	Unique identifier
country	string	Name of the country where company is operational
Domain	string	Information of industry company operation for e.g. real estate

I.2 deal

Column	Data type	Description
id	int	Unique identifier for the deal
company	int	Company reference
user	int	company employee reference
status	boolean	Deal success status. true for won deal, false for lost deal
participant_ count	int	Number of client participants
product_ count	int	Number of products
closed_ date	datetime	Date and time stamp when deal is marked as won or lost
estimated_ close_ date	datetime	Estimated closing date
added_ date	datetime	Timestamp of deal creation
active_days	int	Number of days deal is active

I.2 activity

Column	Data type	Description
id	int	Unique identifier
deal	int	Deal reference
type	string	Type of activity
marked_as_done_date	datetime	Timestamp of activity completion
due_date	datetime	Timestamp of expected completion of activity
added_date	datetime	Timestamp of creation

II. Glossary

event label - Label of a corresponding activity being executed, e.g. the activity of a business process.
event log - Description of an event which has at least two attributes: the label of the event and its timestamp.
event prefix - A partly-complete event trace followed by an event suffix. event suffix A continuation of an event prefix.
event timestamp - Time indicating when the event has been recorded. event trace A sequence of events, ordered by the event timestamp.
process discovery - A set of techniques that construct a representation of an organisation's current business processes.
process mining - A family of techniques to support the analysis of operational processes based on event logs.
schema - Description of the structure and organization of data in a database system.
unicorn - A company valued at $1 billion or higher.

III. Acronyms

BPM - Business Process Monitoring
PPM - Predictive Business Process Monitoring
PrBPM - Prescriptive Business Process Monitoring
KPIs - Key Performance Indicators
CRM - Customer Relationship Management
DW - Data Warehouse
LSTM - Long-Short Term Memory
SaaS - Software as a Service
VPN - Virtual Private Network
S3 - Simple Storage Service

III. Licence

Non-exclusive licence to reproduce thesis and make thesis public

I, **Amey Chnadrakant Darekar**,

1. herewith grant the University of Tartu a free permit (non-exclusive licence) to

   reproduce, for the purpose of preservation, including for adding to the DSpace digital archives until the expiry of the term of copyright,

   Predicting Next Best Action to improve sales metrics for Pipedrive Customers,

   supervised by Mahmoud Shoush.

2. I grant the University of Tartu a permit to make the work specified in p. 1 available to the public via the web environment of the University of Tartu, including via the DSpace digital archives, under the Creative Commons licence CC BY NC ND 3.0, which allows, by giving appropriate credit to the author, to reproduce, distribute the work and communicate it to the public, and prohibits the creation of derivative works and any commercial use of the work until the expiry of the term of copyright.

3. I am aware of the fact that the author retains the rights specified in p. 1 and 2.

4. I certify that granting the non-exclusive licence does not infringe other persons' intellectual property rights or rights arising from the personal data protection legislation.

Amey Chnadrakant Darekar ***08/08/2023***