IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

Impact Factor: 6.252

Volume 3, Issue 1, January 2023

From Business Objectives to Analytics and Machine Learning Solutions: A Framework for Conceptual Modeling

Bhagyashree Pandurang Gadekar[1] and Dr. Tryambak Hiwarkar[2]

Research Scholar, Department of Computer Science[1]

Professor, Department of Computer Science[2]

Sardar Patel University, Balaghat, MP, India

Abstract: *Analytical methods and machine learning are progressively being incorporated into all kinds of information systems. Despite the excitement around these technologies, contemporary firms nonetheless have trouble utilizing them to fully use their data and solve the company's challenges. Businesses must deal with a variety of challenges while developing business analytics and machine learning solutions, including requirements elicitation, design, development, and implementation. Although conceptual modelling and requirements engineering approaches to the process are important and relevant, little study has been done in this area. In this paper a conceptual modelling framework for business analytics and machine learning solutions that is shown and evaluated. The framework consists of instantiations, meta-models, techniques, design patterns and catalogues, rules, and recommendations. It is made up of three modelling perspectives that each reflect a distinct aspect of a solution or the perspective of a different role in the creation of such systems. Through the capture of stakeholders, strategic goals, choices, questions, and necessary insights, the Business View aids in the elicitation of business analytical needs. The Analytics Design View, which largely focuses on machine learning solutions, aids in the design of the solution by collecting algorithms, metrics, and quality criteria.*

Keywords: Machine Learning

I. INTRODUCTION

Analytical components are becoming more prevalent in software-based products, services, and systems. Despite the hype, many firms struggle to apply machine learning and business analytics. Able Company growth is hindered by poor analytics implementation.

Developing usable business analytics solutions requires understanding the limits of analytical methodologies like machine learning algorithms. One must first establish a business case and then turn it into analytical issues. This includes data pretreatment and feature selection, algorithm selection and trade-off analysis, integrating machine learning models to operational processes, and aligning found applications with business strategy. Analytical and machine learning-savvy executives and stakeholders are needed to tackle these issues.

Prerequisites Business analytics elicitation is difficult. Most initiatives start with unclear and insufficient analytics needs. Stakeholders may know their strategic goals, such as improving marketing campaigns or decreasing inventory levels, but they may not understand how analytical approaches might help them achieve them. Stakeholders and data scientists (those with integrated abilities in machine learning, statistics, databases, and optimization have a conceptual mismatch. This gap complicates the issue.

However, linking analytics to company strategy is necessary to achieve benefit. Failure to align may lead to inaccurate assumptions about how analytics contribute to corporate strategy, lack of leadership support, and failed analytics project execution. To achieve such alignment, the organization must establish its analytics project goals, how to distribute resources, and which data assets to focus on [60]. All businesses should priorities finding, justifying, and proving the need for analytics. This aim requires discovering corporate objectives and translating them into analytics goals.

DOI: 10.48175/IJARSCT-7876

1.1 Objectives:

1. Formally sound and application domain-abstract procedures are the key objective.
2. This helps create productive and successful information systems.
3. Conceptual modelling, a key part of requirements engineering, may help plan, create, and deploy business analytics system implementations.

II. LITERATURE REVIEW

Several scholarly works have helped corporate analytics and machine learning advancement. Business analytics is diverse, as is our research.

Theories, concepts, artefacts, and contributions comprise the thesis. In this chapter, we evaluate existing research in each domain, offer context for our work, and underline how our contributions to the literature are both comparable to and unique from others.

Yu-hua et al. [1] 1. Create a data mining ontology to represent user application requirements, offered solutions, and data mining technique resources. This ontology is part of the Universal Knowledge Grid, a framework for grid-based, networked knowledge discovery systems. This ontology focuses on Function, Algorithm, and Solution. One or more Algorithms perform an Application task, a subset of the Application domain. They created data mining use cases by showing how the ontology may be applied to money laundering.

Bernstein et al. [2] built an Intelligent Discovery Assistant (IDA) prototype to help choose data mining methods for a given challenge. The prototype searches the design space for appropriate data mining operator sequences and rates them based on speed and accuracy. The IDA uses a data mining operator ontology for AI planning (algorithms). The ontology classifies data mining techniques into pre-processing, induction, and post-processing. The ontology records each operator's input, output, precondition, and impacts on data and data mining. Kalousis et al. [98] offer a meta-learner to select the best data mining operators.

Choinski and Chudziak [18] the Ontological Learning Assistant, a data mining ontology for knowledge finding (OLA). This ontology captures data mining concepts and represents pre- and post-processing tasks by integrating a metamodel with OLA processes. Both models share classes and relations. The OLA platform can analyses customer turnover.Ontology analysis inspired a business analytics solution metamodel. Ontologies cannot model and analyse business analytics system designs and requirements. Stakeholders, decision activities, and analytic queries are not captured by these ontologies. These ontologies ignore quality standards and their influences.

III. ILLUSTRATIONS OF THE FRAMEWORK

3.1 Descriptive Cases

Two instances follow. A data scientist, conceptual modeller, and requirements engineer assessed these situations to meet specified goals. This chapter's models are based on two key sources: (1) a treasure of business analytics case studies and white paper materials collected from the Internet, and (2) the author's and participants' real-world data mining efforts in both domains. Models can be clarified with assumptions.

Case-1: A Shopping Mobile App

In the first instance, a firm lets clients buy many things straight from the app. Rewarding loyal consumers helps the firm grow and increase profits. Marketers and loyalty programme managers want to use machine learning to solve business challenges. Company databases hold user profiles, app use, and e-commerce histories.

Case-2: A Grocery Retailer.

Second scenario: food distribution center and shop chain. Online grocery bargains, in-store experience, and logistics and operating costs are the company's aims. Modern business analytics and data sets help businesses achieve their aims. Loyalty cards track customer behaviour. It now collects sensor data from store entrances and metro area populations.

IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

Volume 3, Issue 1, January 2023

Impact Factor: 6.252

IV. CONCEPTUAL MODELING IN BUSINESS ANALYTICS—WHY?

We provide numerous ways the framework can help business analytics solution analysis and creation here. These demonstrate the framework's uses.

4.1 Eliciting Needs for Analytical Tasks

It's hard. Business analytics systems are hard to build [99]. The big gap between corporate stakeholders and analytics specialists is the main culprit. Machine learning and analytical methods, technology, and applications are rapidly expanding the chasm.

Business analytics unfamiliarity hinders system development [115]. Stakeholders recognise the importance and usability of analytics systems in many business settings, but they frequently lack a clear understanding of the analytical skills needed and where they are most effective.

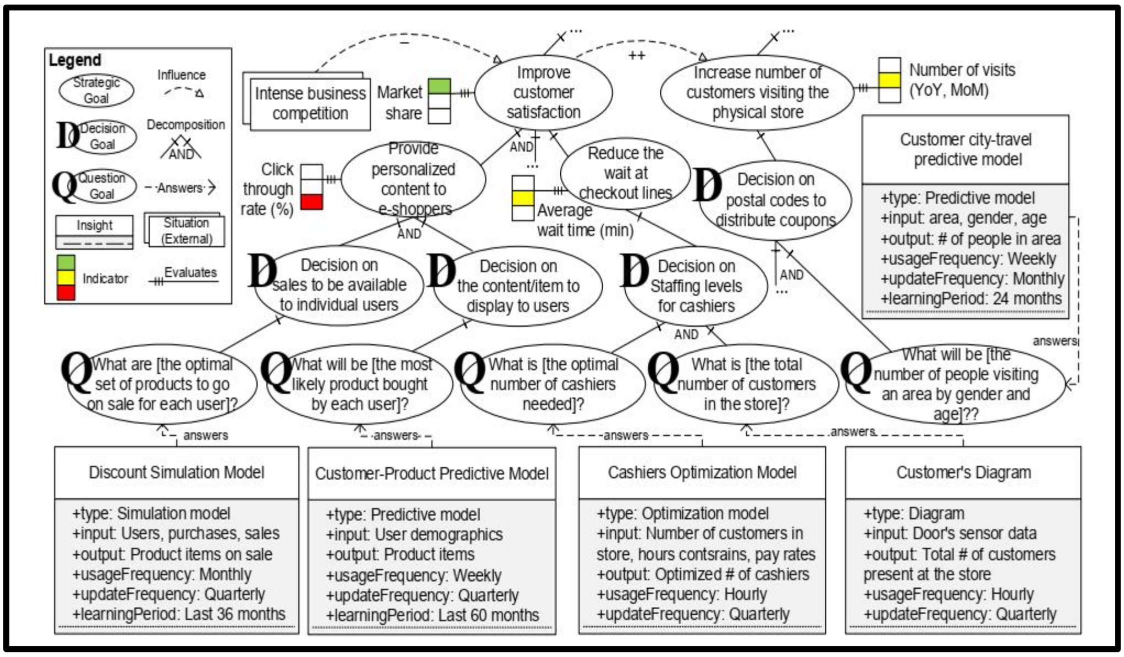

Figure 1: The Grocery Store's Partial Business View Shown

The chasm expands as analytical algorithms, technology, and applications advance.

Business analytics approach unfamiliarity hinders system development, according to research [115]. Stakeholders appreciate the importance and usability of analytics systems in many business contexts, but they frequently lack a clear understanding of the analytical skills needed and where they are most beneficial.

V. DESCRIPTIVE ANALYTICS NECESSITIES

The analytics team and stakeholders often ask for analytics expectations. Analytics programmers fail because the analytics system or team prioritizes different challenges than stakeholders. Data science projects often ask and test a set of wrong questions to iteratively improve the process and arrive at better questions, insights, and conclusions.

The language of business questions can change the analytics task, including algorithm selection, methodology, and data preparation operations. According to study in [6], key stakeholders typically ask ambiguous questions regarding the project's analytics, leading to uncertainty in the definitions of essential variables. To resolve these uncertainties and construct the right business questions, a lot of work and stakeholder engagement is needed. They prevent misinterpretation of analytics findings, saving time and money.

5.1 The Value of Modeling and How It Can Be Used

Figure 1 Business View model shows that the actor needs know the users to choose user interaction methods (a broad question that includes ambiguities). What's each user group's main online activity? What hampers user engagement?

The User Clustering Model addresses this. This Insight piece uses demographics and click data to create user cohorts that answer "What are each user group's key online activities?"

Condensing business issues into sub-questions helps alleviate early stakeholder uncertainty. Decomposition Links subdivide Question Goals breaks down "Who are the users?" into sub-questions. Type, Topic, Tense, and Frequency determine question goals.

Developers (data scientists) and stakeholders can communicate and clarify demands by specifying Question Goal attributes.

Insights demonstrate Question Goal knowledge. Refining Question Goals into sub-questions and defining Insights during modelling helps simplify analytics, reduce ambiguity, and include stakeholders.

VI. ORIGINATING ANALYTICS EXPLANATION ENTERPRISE

Design, algorithm experimentation, and implementation follow analytics demands. Developing more algorithms. Several algorithms exist for numerical prediction, for example (e.g., linear regression, neural networks, and support vector machine). Algorithm selection affects business analytics solution understandability, scalability, memory, noise tolerance, and missing values.

These quality goals determine system success. Algorithm selection requires balancing numerical metrics. Finding the best trade-off is hard.

Case-1 Analytics Design Center model. Analytics Goal: Predict user churn. That requires classifying user profiles and purchases by the analytics system. SVM, Decision Trees, Naïve Bayes, and Neural Networks categories in the model. Algorithm accuracy and sensitivity are examined. The model shows that system development addresses soft goals like tolerance to missing values and noisy input. The model demonstrates how each method influences measurements (numeric labels) and soft goalsqualitative labels. Neural Networks reduce Softgoal Understandability but boost Sensitivity to 0.75. The model chooses SVM with Gaussian kernel function.

Analytics methods depend on design outcomes. The framework includes Type, Analytics Goal, and Generates connections. Insight kind determines business question output.

The analytics goal suggests applicable algorithms. The Algorithm Catalogue (Chapter 3.4) provides Analytics Goal algorithms, metrics, and softgoals. The catalogue lets the project team create the analytics system. The prediction goal into user profile and purchase classification, which various algorithms can do. Business analytics algorithms are selected. Softgoals and Indicators approximate such needs. Goal-oriented reasoning can compare analytics methods. Consider soft goals, their impact, analytics indications, and prioritization. These characteristics are necessary for soft objectives.

VII. EMERGING PROJECT DESIGNS FOR ANALYTICS EXPLANATIONS

Business analytics tools with machine learning are currently available to many firms. Lack of statistics and machine learning experts may hinder analytics adoption. Machine learning's growth complicates system design. Stakeholders must also comprehend machine learning algorithms and business analytics.

7.1 The Value of Modeling and How It Can Be Used

Classification skills are formalized in the model. k-Nearest Neighbor and Random Forest may achieve Classification Prediction Goals. Perceptron and Back-propagation neural networks are shown. The model also shows Recall and Precision as Algorithm Evaluation Indicators. The model also reflects softgoals like speed of learning. Influence Links from Algorithms to Softgoals teach the model how different algorithms perform on certain features.

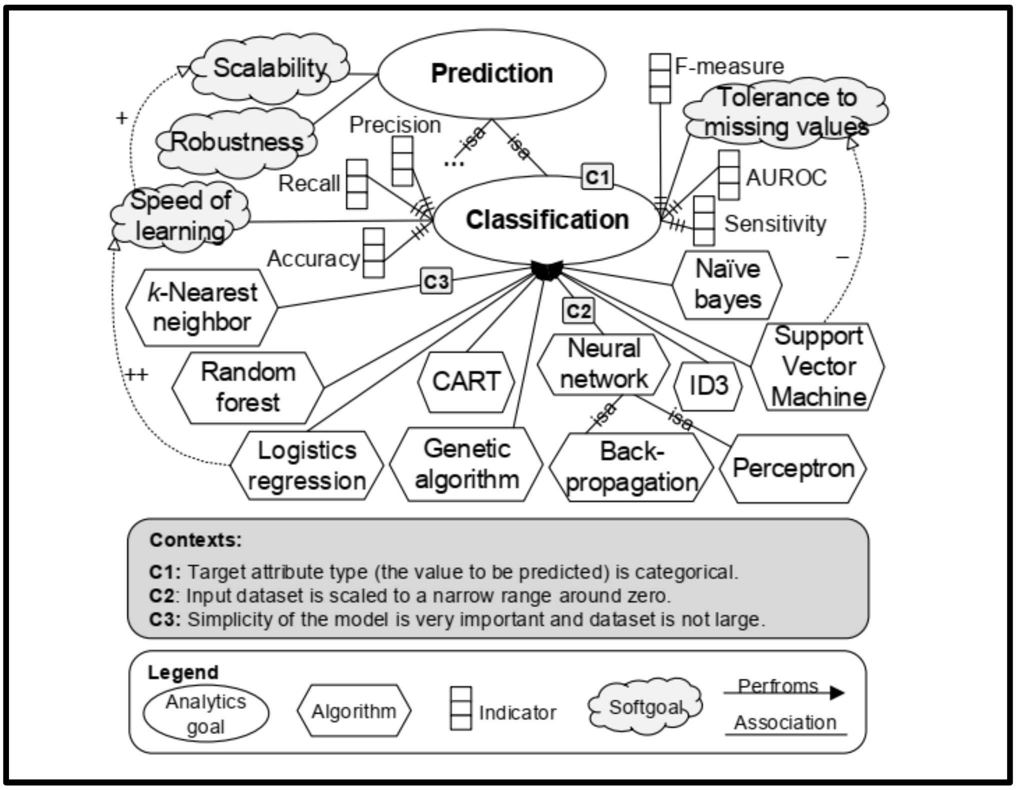

Figure 2: A section from the Catalog of Algorithms

VIII. METHODOLOGY

Next, a sample procedure for developing models in each of the framework's three modelling viewpoints is described. Starting with business goals, the modelling phases include analytical needs, algorithm and solution design, and data preparation. However, bottom-up and hybrid strategies can create and enhance such models. Thus, this method is an example. Additionally, different roles should assist model construction from different viewpoints. Business analysts can build Business View models best. Data scientists build Analytics Design View models (those individuals who devise, execute, or make use of analytical and machine learning strategies). Database administrators and data engineers usually build view models for data preparation (who have a solid understanding of existing data assets, database design and queries in the business domain).

In real-world projects, the problem and project structure might cause roles to vary and overlap. The modeller or analyst must have access to the right stakeholders for modelling and elicitation.

8.1 Sample Methodology Example

A retail mobile app firm's process. 2. Business View models are created from Strategic Goals, Indicators, Situations, and Influences. Strategic Goals in Figure 2 include Improve client retention and Achieve high performance through email marketing. CTR and CR are indicators. Also, low consumer switching costs.

Modeling continues with Decision and Question Goals. Figure 3 shows email decision goals. The actor sending emails to target individuals must choose material to work well. Each user group's most relevant goods? Question Goal. It shows that email content decision-makers must know which commodities are most relevant to each user group cluster.

Answers relate Insight sections to Question Goals. Figure 3 shows User-Product Association Rule Model. Logical principles answer "What are the most relevant elements for each user group?" (e.g., Canadian users with an age between x and y are likely to buy product z. This Insight requires User demographics to identify Products as solutions. 60-month data provides weekly insights. Figure 3 depicts Business View modelling. Analytics Objectives Analytics Design View models incorporate insight features. Classifying user profiles and transactions to predict user churn is shown in Figure 3 (middle). User purchase patterns are descriptive analytics goals.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 3, Issue 1, January 2023
Impact Factor: 6.252

Modeling algorithms, comparison criteria, and performance monitoring continues. Algorithms, Indicators, and Softgoals model this. Pattern-finding algorithms Apriori, ECLAT, and FP Growth. Accuracy and Sensitivity are indicators; Speed of Learning and Tolerance to Missing Values are soft goals.

Define how algorithms effect algorithm selection criteria after modelling. Algorithm-Soft goal-Indicator influence links are created and labelled. This algorithm will hurt softgoal achievement due to the Apriori-Speed of Learning Influence Link. Trials show that FP-Growth and the Indicator % of duplicate rules will produce 0.17 for that indicator. Figure 3 shows each Analytics Design View modelling notion.

Data View models start with understanding data tables, attributes, and linkages. Figure 3 (bottom) illustrates user demographics. This is followed by declaring the dataset(s) prepared for algorithmic analysis. Demographic Product and Churn Variable data tables link to the previous view.

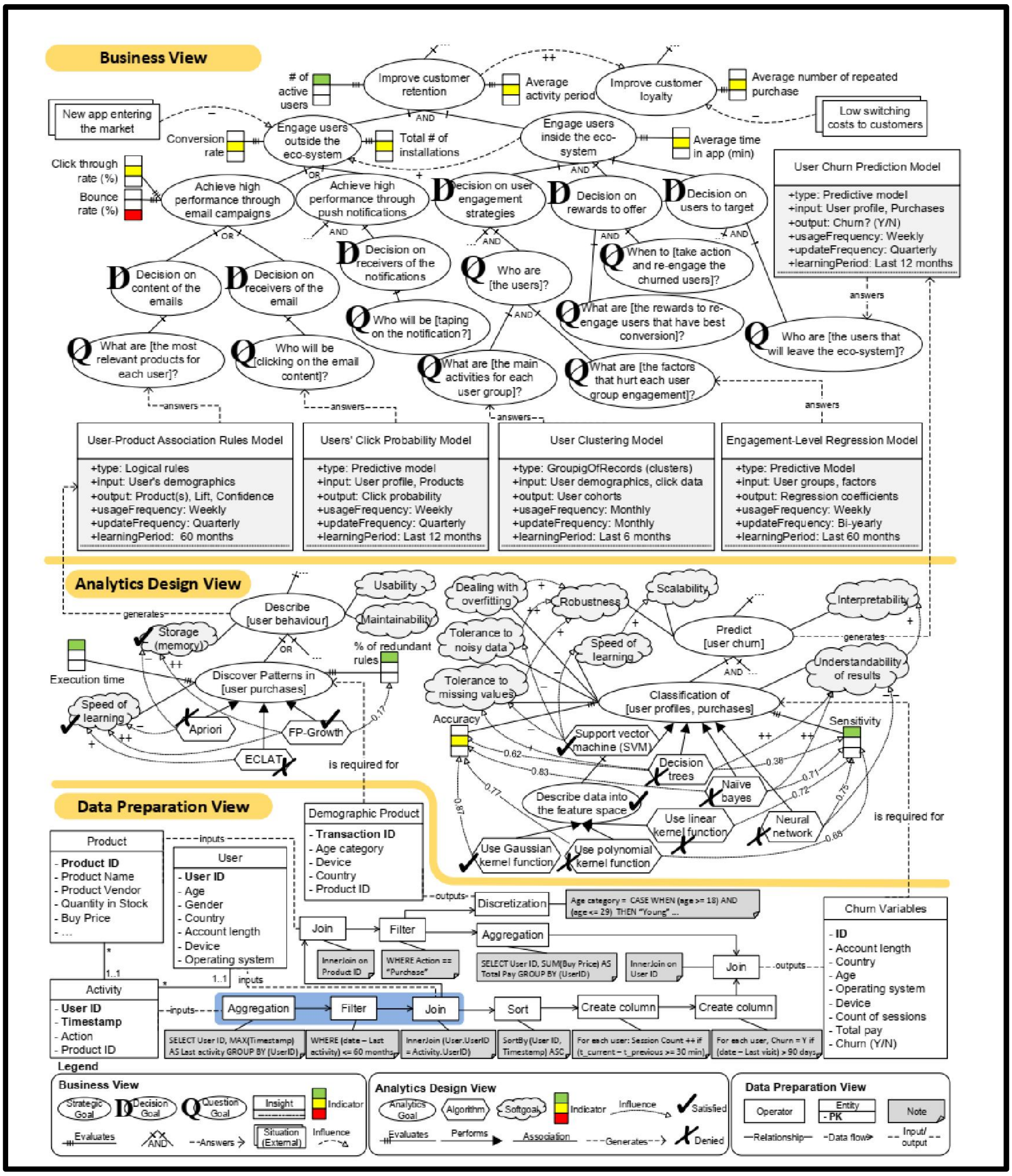

Figure 3: Three shopping mobile app modelling perspectives

IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

Volume 3, Issue 1, January 2023

Impact Factor: 6.252

Modeling determines how to extract and create data tables. Figure 3 shows a Data Reduction operation in the blue-shaded Data Preparation View. The algorithm eliminates five-year non-shoppers. Create column and Join operate. Data flows indicate operator sequence and dependency.

Using Note components, modellers may describe each Operator. Note A new data column for each user is Y if (date { Last visit) > 90 days using a Create column operator. Figure 5.1's bottom illustrates more Data Preparation View ideas.

IX. STRATEGIES

The framework includes instructions. The standards make models easier to use, correct, and comprehend in all three modelling perspectives and make the framework more consistent and successful.

In addition to the observations of a professional who used the framework these guidelines were based on two other sources of information: 1) lessons learned from a project where the framework was tested and models were discussed with real business stakeholders; and 2) the author's experience with goal-oriented modelling techniques, supplemented with benchmarks from existing goal-oriented catalogues (s). Elicitation, Syntax, and Naming Guidelines categories the rules.

X. PATTERN-SOLVING TECHNIQUES FOR MACHINE LEARNING

Design patterns increase design effectiveness and speed up development in software engineering. Design patterns facilitate software development and developer communication by giving well-tested answers to common design difficulties.

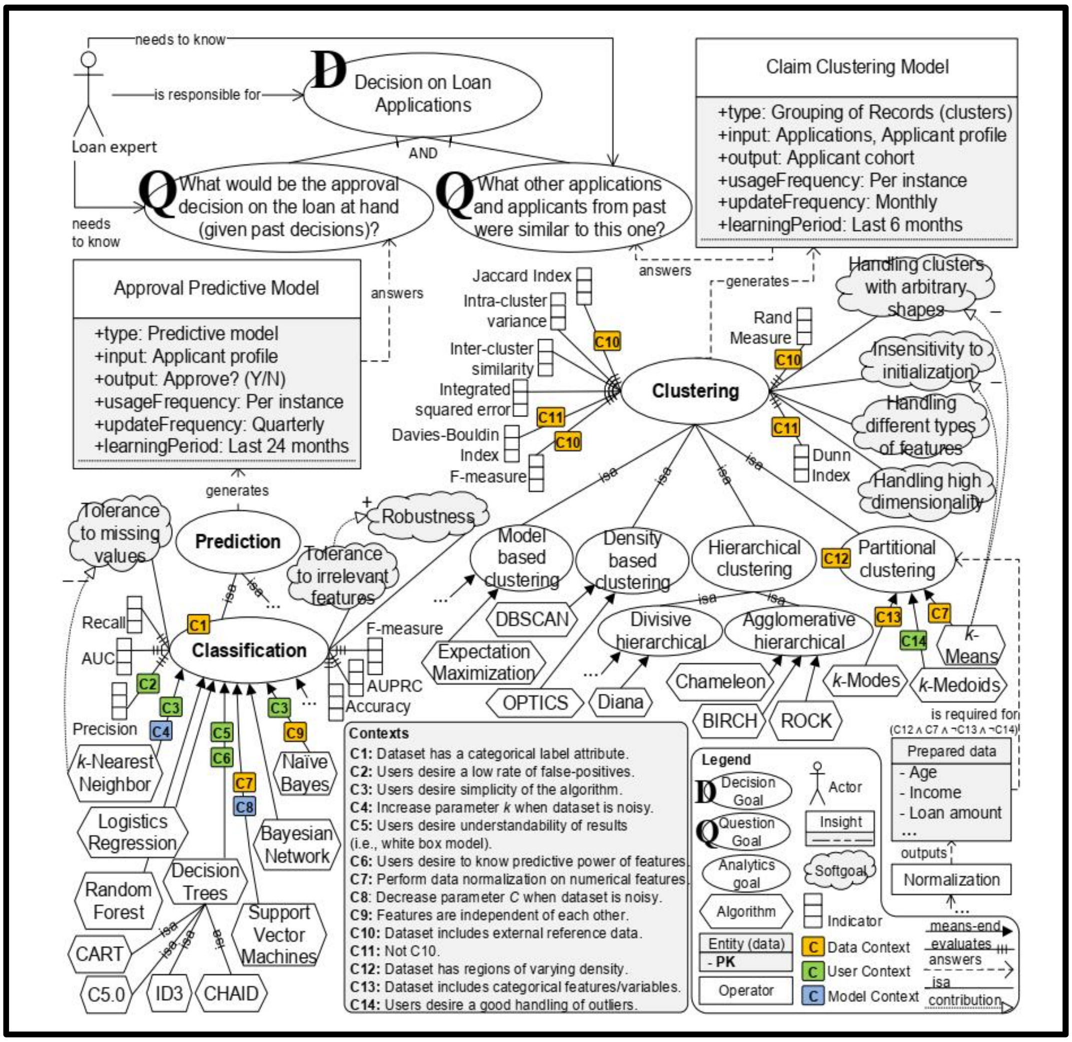

Figure 4: Loan approval solution part. To simplify the figure, Contexts and Contribution Links are not shown.

Covered met models, catalogues, sample techniques, and recommendations. This chapter adds solution patterns to the metamodel to define successful machine learning architectures for business analytics problems. Business analytics solutions can use machine learning and other analytical methods. Business analytics may yield outcomes without machine learning.

Stakeholders, decision activities, business concerns, machine learning algorithms, measurements and parameters, contextual information, quality requirements, datasets, and data preparation procedures comprise a solution pattern. The conceptual model shows general machine learning solutions for business settings.

Solution patterns organize, store, and share machine learning solution design information.

Solution patterns provide reusable answers to machine learning problem-solving issues.

This chapter's solution patterns differ from software design patterns. Design patterns are less problem-specific. These are theoretically closest to a workable solution and should be adopted first.

XI. METAMODEL FOR RESOLUTION DESIGNS

UML class diagram of modelling ideas and semantics. Solution patterns' metamodel incorporates additional components.

Actors set goals. An actor's decision goal is to choose. Decision Goals need actors to answer Question Goals. Question Goals give Actors decision-making information. An Insight is the output of a machine learning solution (e.g., a trained and tested Predictive Model) that answers a Question Goal. Business View metamodel structures reflect the machine learning solution's business issue. Why Actor requires Insight (Decision and Question goals).

Analytics Goals, like Prediction Goals, aim to get data insights. Algorithms achieve Analytics Goals. Algorithms offer a Means-End approach to Analytics Goals. Softgoals and indicators monitor algorithm performance. Analytics View met model describes the business problem's machine learning solution. They demonstrate how to compare and select machine learning algorithms.

Dimensionality Reduction prepares data. Data Flows link operators. Attributed datasets are entities. The Data Preparation View of the metamodeldescribes data preparation operations for machine learning algorithms.

The three views were illustrated in previous chapters. Solution-pattern-specific modelling is covered here. Contexts.

Verification differentiates three cases. Solution users must validate contexts. Data Dataset-verifiable contexts (e.g., size, feature distributions, types). Data transformation and preparation tips. Parameters and algorithm settings check model contexts. They propose experiment and solution sets. This distinction shows what data and user circumstances a solution design including algorithm choice and data preparation methods is appropriate for and how parameters should be specified. Table 1 shows EBNF context expression.

XII. CONCLUSION

The design and development of machine learning solutions may be enhanced by solution patterns.

The cost and time of machine learning development may be reduced by using solution patterns. Instead of starting from scratch, solution patterns allowed the project team to use tried-and-true insurance claim fraud detection techniques and methodologies. Early project study and experimentation were constrained by this. Patterns also guided programmers to libraries and implementations of machine learning algorithms, streamlining code.

(c) Patterns limit the solution space based on the priorities of the modeller and the end user. The fraud detection trend suggests neural network approaches if redundant features are crucial.

(d) Visual patterns facilitated the design of machine learning. The development team was able to comprehend and use the patterns after receiving a quick introduction to goal-oriented modelling languages.

Proposed improvements had to be implemented for the prototype. In the course of implementation, system-verified Softgoals conflicting contexts Systems are required in these situations. Criteria for developing new and extending existing patterns are needed for extensibility. The patterns in the current graphical format can be retrieved, connected to machine learning libraries, and linked using a structured template or domain-specific language.

REFERENCES

[1]. Li Yu-hua, Lu Zheng-ding, Sun Xiao-lin, Wen Kun-mei, and Li Rui-xuan. Data mining ontology development for high user usability. *Wuhan University Journal of Natural Sciences*, 11(1):51{56, 2006.

[2]. Abrahan Bernstein, Foster Provost, and Shawndra Hill. Toward Intelligent Assistance for a Data Mining Process: an Ontology-Based Approach for Cost-Sensitive Classification. *Knowledge and Data Engineering, IEEE Transactions on*, 17(4):503{ 518, April 2005.

[3]. Marcin Choinski and Jaroslaw A. Chudziak. Ontological Learning Assistant for Knowledge Discovery and Data Mining. In *Computer Science and Information Technology, 2009. IMCSIT '09. International Multiconference on*, pages 147{155, Oct 2009.

[4]. Software Engineering for Machine Learning Applications (SEMLA). https://semla.polymtl.ca/. Accessed: 2020-03-07

[5]. Ahmed Abbasi, SuprateekSarker, and Roger HL Chiang. Big Data Research in Information Systems: Toward an Inclusive Research Agenda. *Journal of the Association for Information Systems*, 17(2):1{32, 2016.

[6]. Rama Akkiraju, Vibha Sinha, Anbang Xu, Jalal Mahmud, PritamGundecha, Zhe Liu, Xiaotong Liu, and John Schumacher. Characterizing machine learning process: A maturity framework. *arXiv preprint arXiv:1811.04871*, 2018.

[7]. LianjunAn, JJ Jeng, Markus Ettl, and J-Y Chung. A system dynamics framework for sense-and-respond systems. In *E-Commerce Technology for Dynamic EBusiness, 2004. IEEE International Conference on*, pages 6{13. IEEE, 2004.

[8]. Daniele Barone, Lei Jiang, Daniel Amyot, and John Mylopoulos. Composite Indicators for Business Intelligence. In *International Conference on Conceptual Modeling*, pages 448{458. Springer, 2011.

[9]. Daniele Barone, Eric Yu, Jihyun Won, Lei Jiang, and John Mylopoulos. Enterprise modeling for business intelligence. In Patrick Bommel, Stijn Hoppenbrouwers, Sietse Overbeek, Erik Proper, and Joseph Barjis, editors, *The Practice of Enterprise Modeling*, volume 68 of *Lecture Notes in Business Information Processing*, pages 31{45. Springer Berlin Heidelberg, 2010

[10]. Thomas Binder, Andreas Vox, Salim Belyazid, Hordur Haraldsson, and Mats Svensson. Developing system dynamics models from causal loop diagrams. In *Proceedings of the 22nd International Conference of the System Dynamic Society*, pages 1{21.Citeseer, 2004