

DATA-DRIVEN HYPOTHESIS GENERATION AND DEDUCTION FRAMEWORK FOR BUSINESS INTELLIGENCE

Jyothika A R

PG, Student

Dept. of MCA

The Oxford College of Engineering,
Bommanahalli, Bengaluru- 560068
jyothikaarmca2025@gmail.com

Mridula Shukla

Associate Professor

Dept. of MCA

The Oxford College of Engineering,
Bommanahalli, Bengaluru- 560068
mridulatewari@theoxford.edu

ABSTRACT

The increasing exponentially pace of information in current business environments has set the demand of Business Intelligence (BI) beyond the scope of descriptive analysis and the standard reporting. The current BI tools are useful in providing adequate information on the past and the present trends but fail to enable real-time decision-making procedures due to the inability to generate and test hypotheses. In order to fill this gap, the proposed research proposes a Data-Driven Hypothesis Generation and Deduction Framework for Business Intelligence. The framework utilizes techniques of the big-data analytics, statistical inference, and machine learning to support the automatic creation of potential hypotheses and their testing by means of deductive reasoning methods.

users to recognize suitable medical care for further evaluation of their health.

KEYWORDS: *Business Intelligence Systems, Data-Centric Framework, Automated Hypothesis Generation, Logical Deductive Technology, High End Data Analytics, AI & Learning, Inferential Statistics, Intelligent Decision Backing, Dynamic Decision-Making, Forecasting Analytics*

INTRODUCTION

In today businesses, data has become a very valuable asset to businesses. The increasing exponential growth in the information through digital transactions, social media and IoT sensors is both a great opportunity and poses a challenge in supporting decision-making. Valuable as they are in terms of explaining past and current trends, these systems are unlikely

to be able to support real-time, proactive, and predictive decision-making.

A major handicap is that current BI tools cannot automatically generate, and test hypotheses that drive strategic and operational decision-making. This hand-cranked process of assumptions is time consuming and subjective. A hypothesis generation and deduction environment, which is driven by data in order to convert data to form meaningful conclusions, is therefore proposed. The system follows logics of big data analysis, machine learning, and statistical inference to automatically tease out a hypothesis, which can be tested and confirmed, using the logics of deductive reasoning and inference methods.

LITERATURE SURVEY

The world of Business Intelligence (BI) has taken a gigantic step towards descriptive reporting to advanced analytics guiding decisions in business companies. Traditional BI instruments predominately provide information on the past and present trends and are ineffective at supporting in-time, proactive, and predictive decision-making. A big disadvantage is that it is subject to time-consuming and biased manual hypothesis development. Recent studies have centered on the efficacy of automated hypothesis-generation, such machine learning, statistical inference or data mining, emphasis or pattern

discovery. Such approaches can help an organization determine hidden patterns, relationships and potential cause and effect within large and heterogenous data.

Some efforts have been made to redesign predictive analytics and real-time monitoring to integrate it with BI systems to provide dynamic decision support. Few frameworks, however, can combine automated hypothesis generation and rigorous validation with velnet-like BI fully integrated into a system into one stack.

The literatures reveal a clear need; the need of having a data-driven hypothesis formation and inference method that transforms the raw data into constructive valid inferences. With enrichment of BI systems in companies, this framework enables business to enable strategic, proactive and content-based business decisions, to realize continuity value in their tactical processes through new sources of data.

EXISTING WORK

Living with latent allotments in the current systems is rather an ad hoc and illogical practice. Because of using hypothetical activity and categorization of strategy into broken-down categories, people cannot easily feel that this or that task is associated with a more strategic goal. This disjointed structure complicates multi-level planning at the

expense of more money, time and reduced efficiency in operations.

The collaborative process of digitalization is limited as well because the provided frameworks do not imply centralized collaboration among different groups. The amount of time it takes to come up with the decision is increased because of the ad hoc cooperation, the degree of transparency is decreased, and there are no possibilities of coordination in real-time situations because there is no united frame of reference provided.

Repeat and generation of information/printer of paper requires much manual use and this is time-consuming, involves many efforts, creates inaccuracies and inefficiencies. The redrafting of reports in regards to the hypothesis planning is cumbersome and there is lack of coordination of the pieces in a kinship fashion without automated functions

The part-base embedding or the process tracing is weak where it is not easy to trace the dependencies, design comprehension or convenience to make changes. Process flow and visualisation tools are not combined and this hinders the efforts of strategic control by the users.

PROPOSED SYSTEM

The The high co-operation and understanding of multi-dimensional phenomena occurring

with the support of the under consideration system can be facilitated by the proposed single scheme through which the tasks and remote classification can be addressed. The openness of the system allows users to create an enhanced flow of the process to include resource management and prediction as well as viewing of the strategic objectives.

An institutionalized and transparent process modelling that is grounded on a formal process description and visualisation, will make more advanced strategies, more easily understandable and permit more sophisticated decisions to be made. The user will be in a position to monitor workflows simplifying everything being logically logical.

In addition to the ability to parallelly process multiple hypotheses, the system has many utilities, which are provided in diverse contexts, and reduce the bottlenecks, and maximize the overall efficiency. Multi-task management is the process in which ensures the flawless execution of complex scenarios without multi-tasking and distractions.

The essence digitalization makes it possible to collaborate in a timely manner sets, provide counters to multi-channel replications of information and concerted data exchange.

categories and built-in objects minimize the preparation of documents across formats and relational outlining and dependency tracking

through automation ensure the proper and synchronized updating of documents. The component-based integration allows viewing of information retrieval, process flow visualization, and reporting to third party platforms.

On the whole, the presented system can simplify the planning of strategies, make cooperation more convenient, enhance efficiency and ensure the proper, consistent, and prompt piece of information. Its features of automated processes, simultaneous multi-tasking and collaborative communication removes the limitations of current systems and makes the process of evidence-based making decisions possible.

METHODOLOGY

The proposed framework is one that applies systematic approach to the mechanization of the creation, deduction and unification of hypotheses in Business Intelligence (BI) systems. This is followed by data preprocessing and loading where a varied mass of data sources such as transaction records, user activities, and external data are preprocessed and cleansed, be normalized, and unified to maintain a uniformity of data.

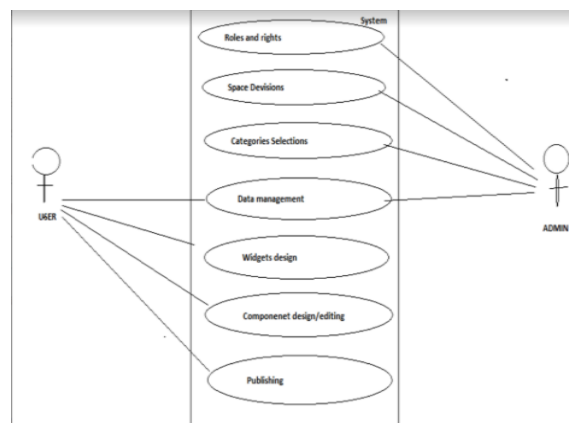


Fig.1.Block Diagram

Second is the hypothesis generation module, which applies machine learning, pattern mining, and statistical inference to the automatic generation of possible hypotheses. Unstructured data sources, like text, can also be analysed through Natural language processing (NLP) as a method of deriving insights.

Task management and process flow module ensures standardized visualization and component-based embedding so that the users can trace the dependencies, workflow mapping and parallel running of the hypothesis activities in a resourceful way. Digitalization and collaboration efforts are centralized thus allowing real-time updates, multi-user access and multi-channel replication of information within a team

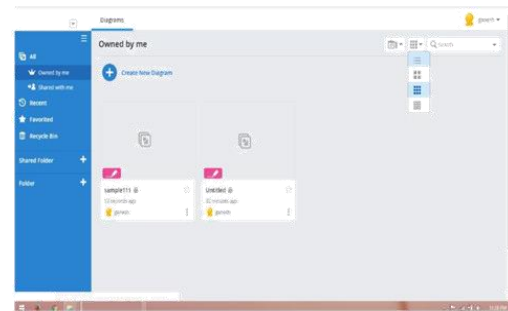
EXPERIMENTAL RESULTS

A login page provides a secure way to access a system, web site, or application with some credentials such as a username/email and a password. Upon authenticating, one is redirected to a dashboard that contains tools to create and manage diagrams and having navigation panels that provide direct access to categories like: All, Owned by Me, Shared with Me, Recent, Favorited and Recycle bin. User is able to create new projects using a button titled Create New Diagram and switch back and forth between grid or list views, to simplify organization.

Wireframes present a grey shade visual layout framework of a web page or application which puts more emphasis on hierarchy and functionality than appearance.

Templates are readymade, reusable document, presentation, diagram and web site templates that are time saving and ensure uniformity and professionalism. Business presentations employ the use of slides, charts, graphics to promote an abridged presentation of information informing decision-making and involvement.

Customization allows the user to alter settings, layout or content to fit his/her personal needs, which results in satisfaction and product ownership.



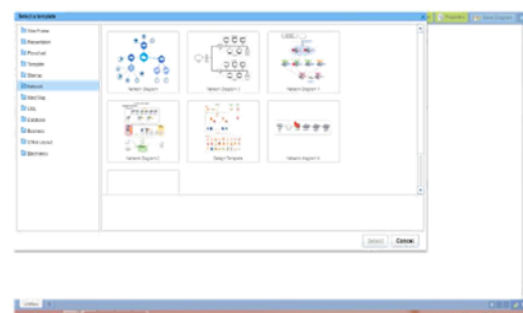
Home page

Fig.1.Home page



Sharing properties shown

Fig 2. Sharing properties shown



Templates provided for reusability

Fig 3. Template

CONCLUSION

To ensure proper integration of planning perspectives, sizeable plans projects are integrated into the process flow and accompanied by the design of a node structure of design. The reason is that each strategy is a distinct element of the entire system and, therefore, the project goals can be aligned properly. By use of in-depth analysis, all the components are clearly charted to ease consistency and coherence.

Data presentation and evaluation is also effective, as it involves the integration of predicted data in into the design requirements that have been set by a user. This makes the interpretation of data very clear and decision-making to be undertaken by the stakeholders.

Aggregated platforms and identification systems ensure precision, neutrality and concise representation. The real-time feature allows real-time monitoring the updates and revisions and ensuring transparency within the complex processes.

This systematic definition of projections and organization of strategies enables more intricate procedures to be easier, more transparent, and manageable. In such an organized process, strategic goals are reconciled with operational delivery, project management is simplified, and effective,

evidence-based decision-making is conducted throughout the entire duration of the project.

REFERENCES

- [1] "Business Intelligence and Big Data", Drivers of Organizational Success by Celina M. Olszak
- [2] "MongoDB 8.0 in Action", Third Edition by Arek Borucki
- [3] Web Programming with HTML, CSS, Bootstrap, JavaScript, jQuery, PHP, and MySQL" by Larry Sanchez
- [4] "Learning Bootstrap" by None Shenoy
- [5] "Eloquent JavaScript" by Marijn Haverbeke