A REPORT ON

Data Science and Big Data Analytics Mini Project

on

Case study on Global Innovation Network and Analytics

SUBMITTED TO THE SAVITRIBAI PHULE PUNE UNIVERSITY, PUNE IN THE PARTIAL FULFILLMENT FOR THE AWARD OF

THE DEGREE OF BACHELOR OF ENGINEERING IN COMPUTER ENGINEERING

BY

Ms. Aarti Swami 3202065

Under The Guidance

Mr. Sachin Dighe

DEPARTMENT OF COMPUTER ENGINEERING

Sinhgad Institute of Technology & Science Narhe, Pune - 411041

Sinhgad Institutes

Department of Computer Engineering, Sinhgad Institute of Technology and Science, Narhe

CERTIFICATE

This is to certify that the report entitled

"Case study on Global Innovation Network and Analytics"

Submitted by

Ms. Aarti Swami 3202065

is a bonafide work carried out by them under the supervision of **Mr. Sachin Dighe** and is approved for the partial fulfillment of the requirement of **Data Science and Big Data Analytics Laboratory** course in Third Year Computer Engineering, in the academic year 2022-2023 prescribed by Savitribai Phule Pune University, Pune .

Mr. Sachin Dighe	Dr. G. S. Navale	Dr. S. D. Markande
Guide	**Head of Department**	**Principal**

SINHGAD INSTITUTE OF TECHNOLOGY AND SCIENCE, NARHE, PUNE– 411040031

Place: Pune

Date :18/05/2023

Acknowledgement

I take this opportunity to acknowledge each and every one who contributed towards my work. I express my sincere gratitude towards guide **Mr. Sachin Dighe** Assistant Professor at Sinhgad Institute of Technology and Science, Narhe, Pune, for her valuable inputs , guidance and support throughout the course.

I wish to express my thanks to **Dr. G. S. Navale,** Head of Computer Engineering Department, Sinhgad Institute of Technology and Science, Narhe for giving me all the help and important suggestions all over the Work. I thank all the teaching staff members, for their indispensable support and priceless suggestions.

I also thank my friends and family for their help in collecting data without which their help Database Management Laboratory report have not been completed. At the end my special thanks to **Dr. S.D.Markande,** Principal Sinhgad Institute of Technology and Science, Narhe for providing ambience in the college, which motivate us to work.

Aarti Swami

Contents

Introduction

In today's rapidly evolving and interconnected world, innovation has become a driving force for economic growth, technological advancements, and sustainable development. Organizations across industries are increasingly recognizing the need to understand global innovation trends and leverage them to gain a competitive edge. The Global Innovation Network and Analysis (GINA) is a comprehensive data-driven approach designed to address this need. This case study explores the components of the analytic plan for GINA, highlighting its significance in guiding organizations towards strategic decision-making and fostering collaboration in the global innovation landscape.

The purpose of GINA is to provide organizations with valuable insights into emerging technologies, collaboration opportunities, and innovation hubs worldwide. By analyzing diverse data sources, such as patent records, research publications, funding data, and collaboration networks, GINA uncovers hidden patterns, identifies key players, and predicts future trends. These insights enable organizations to refine their innovation strategies, identify potential partners, and allocate resources effectively.

At its core, GINA seeks to solve the complex business problem of understanding and harnessing global innovation. Innovation, in this context, refers to the creation, adoption, and diffusion of new ideas, technologies, and practices. Organizations face numerous challenges in navigating the dynamic global innovation landscape, such as staying up-to-date with emerging technologies, identifying potential collaborators, and capitalizing on opportunities in different markets. GINA provides a systematic framework to address these challenges and support evidence-based decision-making

Problem Statement

Write a case study on Global Innovation Network and Analysis (GINA). Components of analytic plan are 1. Discovery business problem framed, 2. Data, 3. Model planning analytic technique 4. Results and Key findings.

Description

Phase 1: Discovery

Everything starts with a purpose in mind. In this phase, you will identify the objective of your data and how to achieve it after the data analytics Life Cycle. The purpose of this initial phase is to conduct evaluations and assessments to develop a fundamental hypothesis for resolving any business problems and issues.The first stage entails mapping out the potential use and demand of data, such as where the data is coming from, the story you want your data to portray, and how your business benefits from the incoming information.

As a data analyst, you will need to explore case studies using similar data analytics and, most crucially, examine current company trends. Then you must evaluate all in-house infrastructure and resources, as well as time and technological needs, in order to match the previously acquired data.

Phase 2: Data Preparation

The data preparation and processing phase involves gathering, processing, and purifying the collected data. One of the most important aspects of this step is ensuring that the data you require is available for processing.

The following techniques are used to acquire data:

- Data Acquisition: Accumulate data from external sources.
- Data Entry: Within the organization, creating new data points utilizing digital technologies or manual data input procedures.
- Signal Reception: Information gathering from digital devices such as the Internet of Things and control systems.

An analytical sandbox is required during the data preparation phase of the data analytics Life Cycle. This is a scalable platform used to process data by data analysts and data scientists. The analytical sandbox contains data that has been executed, loaded, and altered. **Phase 3: Model Planning**

After you've defined your business goals and gathered a large amount of data (formatted, unformatted, or semi-formatted), it's time to create a model that uses the data to achieve the goal. Model planning is the name given to this stage of the data analytics process.There are numerous methods for loading data into the system and starting to analyze it:

ETL (Extract, Transform, and Load) converts the information before loading it into a system using a set of business rules.

ELT (Extract, Load, and Transform) loads raw data into the sandbox before transforming it.

ETLT (Extract, Transform, Load, Transform) is a combination of two layers of transformation.

This step also involves teamwork to identify the approaches, techniques, and workflow to be used in the succeeding phase to develop the model. The process of developing a model begins with finding the relationship between data points in order to choose the essential variables and, subsequently, create a suitable model.

Phase 4: Model Building

This stage of the data analytics Life Cycle involves creating data sets for testing, training, and production. The data analytics professionals develop and operate the model they designed in the previous stage with proper effort.They use tools and methods like decision trees, regression techniques (such as logistic regression), and neural networks to create and run the model. The experts also run the model through a trial run to see if it matches the datasets.It assists them in determining whether the tools they now have will be enough to execute the model or if a more robust system is required for it to function successfully.

Phase 5: Operationalize

Recall the objective you set for your company in phase 1. Now is the time to see if the tests you ran in the previous phase-matched those criteria.The communication process begins with cooperation with key stakeholders to decide whether the project's outcomes are successful or not.The project team is responsible for identifying the major conclusions of the analysis, calculating the business value associated with the outcome, and creating a narrative to summarize and communicate the results to stakeholders.

Phase 6: Communicate Results

As your data analytics Life Cycle comes to an end, the final stage is to offer stakeholders a complete report that includes important results, coding, briefings, and technical papers/documents. Furthermore, to assess the effectiveness of the study, the data is transported from the sandbox to a live environment and observed to see if the results match the desired business aim. If the findings meet the objectives, the reports and outcomes are finalized. However, if the conclusion differs from the purpose stated in phase 1. Then You can go back in the data analytics Life Cycle to any of the previous phases to adjust your input and get a different result.

Conclusion

The Global Innovation Network and Analysis (GINA) serves as a valuable tool for organizations seeking to thrive in the ever-evolving global innovation landscape. By framing and addressing the business problem of understanding and leveraging global innovation, GINA provides a systematic approach that supports evidence-based decision-making and fosters collaboration. Through the analysis of diverse data sources, GINA uncovers valuable insights, identifies emerging trends, and maps collaboration opportunities.

The components of the GINA analytic plan, including data collection, model planning, and analytic techniques, work together to generate meaningful results and key findings. These findings enable organizations to refine their innovation strategies, identify potential partners, and allocate resources effectively. By leveraging the power of data analytics, GINA empowers organizations to make informed decisions, stay ahead of emerging technologies, and capitalize on opportunities in different markets.

In conclusion, the Global Innovation Network and Analysis (GINA) represents a crucial tool for organizations seeking to navigate the complexities of the global innovation landscape. By leveraging data analytics and advanced techniques, GINA enables organizations to gain valuable insights, make informed decisions, and foster collaboration in the pursuit of innovation. As organizations embrace GINA and its analytic plan, they position themselves to thrive in the dynamic and interconnected world of global innovation.