MES Wadia College of Engineering Pune-01

Department of Computer Engineering

Name of Student: Khushal Patil	**Class: TE Comp – 2 (R Batch)**
Semester/Year: 6Th / 3rd Year	**Roll No: 64**
Date of Performance:	**Date of Submission:**
Examined By:	**Experiment No:** Case-Study (C-1)

PART: C) ASSIGNMENT NO: 01

Title: Case Study

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 and 4. Results and Key findings.

Introduction

The Global Innovation Network and Analysis (GINA) team is a group of senior technologists located in centers of excellence (COEs) around the world. Their charter is to engage employees across global COEs to drive innovation, research, and university partnerships.

This case study examines how the GINA team applied the Data Analytics Lifecycle to analyze innovation data at EMC. Innovation is a complex and difficult concept to measure, and the team sought to leverage advanced analytical methods to identify key innovators within the company. The approach aimed to facilitate global knowledge sharing, track research progress, and analyze innovation trends using structured and unstructured data.

The study's primary objectives included:

1. Storing formal and informal innovation-related data.
2. Tracking research contributions from global technologists.
3. Mining data for patterns and insights to enhance strategic decision-making.

Phase 1: Discovery

- In the GINA project's discovery phase, the team began identifying data sources.
- The Various Roles are involved in this phase.
    1. Business user, Project Sponsor, Project Manager: - Vice President from Office of CTO
    2. BI analyst person from IT
    3. Data Engineer and Database Administrator – people from IT

4. Data Scientist: - distinguished engineer

- The data for the project fell into two main categories:
    1. Innovation Roadmap.
    2. Data encompassed minutes and notes representing innovation and research activity from around the world.
- The GINA (Hypothesis) can be grouped into two categories: -
    - Descriptive analytics of what is currently happening to spark further creativity, collaboration and asset generation.
    - Predictive analytics to advise executive management of where it should be investing in the future. Global Innovation Network.

Phase 2: Discovery

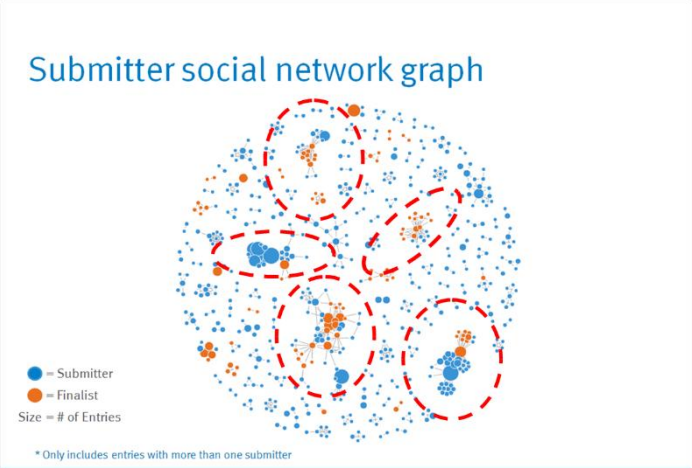
- IT Department to set up a new analytics sandbox to store an experiment on the data.
- The data scientists and data engineers begin to notice that certain data needed conditioning and normalization.
- As the team explored the data, it quickly realized that if it did not have data of sufficient quality or could not get good quality data, it would not be able to perform any subsequent steps in the lifecycle process.
- Important to determine what level of data quality and cleanliness was sufficient for the project being undertaken.

Phase 3: Model Planning

- The team made a decision to initiate a longitudinal study to begin tracking data points over time regarding people developing new intellectual property.
- The parameters related to the scope of the study included the following considerations:
    1. Identify the right milestones to achieve this goal.
    2. Trace how people move ideas from each milestone toward the goal.
    3. Once this is done, trace ideas that die, and trace others that reach the goal. Compare the journeys of ideas that make it and those that do not.
    4. Compare the times and the outcomes using a few different methods (depending on how the data is collected and assembled). These could be as simple as t-tests or perhaps involve different types of classification algorithms.

Phase 4: Model Building

- The GINA team employed several analytical methods. This included work by the data scientist using Natural Language Processing (NLP) techniques on the textual descriptions of the Innovation Roadmap ideas.
- Social network analysis using R and RStudio.



- Fig shows social graphs that portray the relationships between idea submitters within GINA.
    1. Each color represents an innovator from a different country.
    2. The large dots with red circles around them represent hubs. A hub represents a person with high connectivity and a high "betweenness" score.
- The team used Tableau software for data visualization and exploration and used the Pivotal Greenplum database as the main data repository and analytics engine.

Phase 5: Communicate Results

- This project was considered successful in identifying boundary spanners and hidden innovators.
- The GINA project promoted knowledge sharing related to innovation and researchers spanning multiple areas within the company and outside of it.
- GINA also enabled EMC to cultivate additional intellectual property that led to additional research topics and provided opportunities to forge relationships with universities for joint academic research in the fields of Data Science and Big Data.

Phase 6: Operationalization

- Deployment was not really discussed.
- Key findings:
    1. Need more data in the future.

2. Some data were sensitive.
3. A parallel initiative needs to be created to improve basic BI activities.
4. A mechanism is needed to continually reevaluate the model after deployment.

Components of Analytic Plan - GINA Case Study

Components of Analytic Plan	GINA Case Study
Discovery Business Problem Framed	Tracking global knowledge growth, ensuring effective knowledge transfer, and quickly converting it into corporate assets. Executing on these three elements should accelerate innovation.
Initial Hypotheses	An increase in geographic knowledge transfer improves the speed of idea delivery.
Data	Five years of innovation idea submissions and history, six months of textual notes from global innovation and research activities.
Model Planning Analytic Technique	Social network analysis, social graphs, clustering, and regression analysis.
Result and Key Findings	1. Identified hidden, high-value innovators and found ways to share their knowledge. 2. Informed investment decisions in university research projects. 3. Created tools to help submitters improve ideas with idea recommender systems.

CONCLUSION:

GINA (Global Innovation Network and Analysis) is a tool that provides insights and analysis to support innovation and technology development. It enables companies and organizations to stay up-to-date with the latest industry trends and technology advancements, allowing them to make informed decisions and stay competitive in their respective fields. The platform offers a range of features, including technology scouting, competitor analysis, and IP portfolio management, and can be customized to suit the needs of individual companies. Overall, GINA is a valuable resource for companies looking to innovate and stay ahead of the curve in their industries.