Indian School of Business

ISB365

March 27, 2023

Sumanta Singha | Milind Sohoni | Sripad Devalkar | Vijaya Sunder M

ARTIFICIAL INTELLIGENCE FOR IMPROVING THE PROCUREMENT EXPERIENCE OF NON-STOCK ITEMS AT INDIAN RAILWAYS

INTRODUCTION

During the summer of 2021, Sumana G., Chief Technology Officer of South Central Railway, was reviewing the annual productivity reports of field employees. This was an annual exercise that was crucial to central planning as it helped identify potential weaknesses and possibilities for improvement. Sumana knew that evaluating the productivity of store personnel would be the most challenging task because Indian Railways (IR) managed over 280,000 items stocked in 215 depots across the country. While reviewing the time sheets, Sumana quickly realized that field officers were spending a significant time amount of time on materials purchase, especially items purchased locally by field offices. On further inquiry, field officers revealed that retrieving data from the stores database based on item descriptions posed considerable challenges, and in most cases, the search results were not very useful. Sumana was quick to realize that an artificial intelligence (AI)-based search engine could solve this problem.

BACKGROUND

IR was the largest national rail network in Asia, the fourth-largest rail network globally, and the world's second-largest network operated under a single management as of 2021.[1] It had played a significant role in India's economic development for over 160 years since its inception in the 1850s. IR accounted for 75% of public transport and 90% of freight in 2021. With close to 68,000 kilometers of route length, IR carried over eight billion passengers in 2020.[2] The country had over 9,000 freight trains in operation, which amounted to a daily freight transport volume of over three million metric tons as of March 2020.

---

[1] Singh, S. P. (2015, February 16). 18 interesting facts about India Railways. *Business Standard*. Retrieved from https://www.business-standard.com/article/beyond-business/18-interesting-facts-about-india-railways-business-standard-news-115021600404_1.html.

[2] Directorate of Economics and Statistics, Ministry of Railways (Railway Board), Government of India. (2020). *Indian Railways Yearbook 2019-20*. New Delhi: Directorate of Economics and Statistics, Ministry of Railways (Railway Board), Government of India. Retrieved from: https://indianrailways.gov.in/railwayboard/uploads/directorate/stat_econ/Annual-Reports-2019-2020/Year-Book-2019-20-English_Final_Web.pdf

Professor Sumanta Singha, Professor Milind Sohoni, Professor Sripad Devalkar, and Professor Vijaya Sunder M prepared this case solely as a basis for class discussion. This case is not intended to serve as an endorsement, a source of primary data, or an illustration of effective or ineffective management. The authors would like to acknowledge the support provided by G. Sumana, Chief Technology Officer, South Central Railway, in the writing of this case. This case was developed under the aegis of the Centre for Learning and Management Practice, ISB.

IR was divided into 18 zones under the apex management of the Railway Board. To support efficient and punctual operations, the IR zones managed over 0.25 million different materials stocked across 215 depots across the country. The materials were broadly classified into stock and non-stock items based on usage. Stock items such as bogies, wagons, and wheels were primarily strategic and were characterized by regular demand, regular consumption, and regular recoupment. These were largely standardized and thus scheduled for procurement, coordinated by the central stores in each zone. Non-stock items such as seating fabric, cutlery, water geysers, and compressors were relatively small in size, but accounted for more than 70% of the overall inventory. These items did not require regular replenishment and hence were procured only for a particular activity as and when they were required. Non-stock items were purchased locally by field units (depots within each zone) to fulfill short-term, primarily one-off needs. Each item was assigned a unique Price List (PL) number, which was used to determine its category (stock vs. non-stock item), perform item reconciliation, and facilitate reordering.

MATERIAL PROCUREMENT

Similar to most large corporations, IR had to manage a massive inventory to support its pan-India operations. The store manager of South Central Railway understood that procuring the right materials at the right time and at the right price was the key to profitable and sustainable railway operations. The challenge for IR was to streamline the existing procurement process and eventually build a strong database to improve consistency regarding price and quality. Over many decades, IR had managed to build a robust procurement system for stock items that were ordered, negotiated, and procured centrally for its different divisions. However, there were challenges in the procurement of non-stock items that were occasionally required and hence purchased locally by various field units as the need arose. Sumana observed considerable variation in the descriptions of the non-stock items in the database. The nomenclature used by field officers to describe the same item varied, making it harder for them to retrieve past orders using a simple text search, let alone a price comparison. Sumana’s intuition was, in fact, correct: this was a bigger problem across IR and not merely at the South Central Railway, which was one of the 18 zones of IR.

CURRENT PRACTICE

The annual value of store procurement for operations, repairs, maintenance, and construction was about INR 207 billion in 2020, excluding fuel. The procurement and materials management for such voluminous and high-value operations naturally posed considerable challenges for IR. To facilitate this process, IR implemented an integrated monitoring system (IMS), an intranet-based supplier management system. IMS had a search functionality similar to Google search: field officers typed item descriptions to check the availability of similar items in other zones, vendor details, and the rates at which the same/similar items were procured in the past.

It was evident to Sumana that the current keyword-based search system required significant human intervention to identify relevant results from the IMS database. She knew that standardizing how field officers searched IMS would require considerable change management and was therefore a difficult solution to implement. Deming’s famous quote—"94% of problems in business are systems-driven, and only 6% are people-driven"—resonated with her. She decided to take the AI route to enhancing the search experience rather than changing people's behavior (by standardizing the search keywords and insisting that people use them). She consulted Harsha Reddy, a data analytics consultant at the Indian School of Business, to help solve the item-matching problem. The plan was to develop a text-based clustering algorithm that considered not only the lexical similarity between two text strings but also their semantic match.

THE MISMATCH PROBLEM

Harsha realized that an incorrect search string used by field officers ran the risk of either being overly broad (in which case the officer would have to sift through voluminous data to identify relevant items) or too restrictive (in which case some relevant purchases may be missed in the search). Upon further investigation, Harsha found that unless the search string was identical to the description in the IMS database, it was not considered a "match." Thus, the system failed to return a result if the query described items somewhat differently and/or keywords appeared in a different sequence in the query. Therefore, when field officers used search strings that differed from the corresponding descriptions in IMS, they received a message that the particular non-stock item was either not available in the database or was not offered by empaneled vendors. This resulted in purchase officers wasting considerable time and effort making multiple keyword searches to obtain the required data. Despite these efforts, on 15,000 occasions, field officers failed to find the required non-stock item in IMS in 2020. As one of the field officers explained:

> "While placing an order for a grinder, I used the search string "grinder", and IMS showed me 200 search results. It took me more than 30 minutes to go through them and identify the correct choice. In another instance, I had to use a combination of keywords like "grinder", "angle grinder", "bosch", "twisted cup", "spindle", etc., before realizing that we do not have a vendor who provides angle grinders with the required spindle diameter and a generic 3-inch twisted cup brush," said a field officer to Sumana.

Interestingly, Sumana found that the item matching problem occurred not only at South Central Railway but across all IR zones. She was determined to resolve the problem, as the search times lowered the productivity of field officers. Further, as explained in the next section, failed searches resulted in unnecessary vendor onboarding procedures, leading to redundant efforts.

CONSEQUENCES

When field officers failed to find the required non-stock material at IMS, they initiated a vendor empanelment process to onboard a new vendor who could provide the item. The empanelment process commenced with the user department preparing an indent that specified the technical and non-technical specifications of the non-stock requirement. Subsequently, the Stores Department initiated a request for proposal (RFP) based on the requirement. Moreover, many local vendors responded to tenders and submitted proposals. After receiving proposals from various vendors, the purchase officer assessed them by following an approval procedure to onboard the vendor, which took about two weeks. Overall, onboarding a new vendor took anywhere between five to eight weeks.

In many cases, after spending eight weeks on vendor onboarding, the field/zonal officers found that the material for which the vendor onboarding was initiated was available in the IMS database. Thus, the database search failures created confusion and triggered redundant efforts. Further, the inability to compare prices hampered project progress and increased the administrative burden, thereby lowering productivity.

TOWARD AN AI-BASED SOLUTION

Sumana approached Reddy to seek his guidance on improving search efficiency for non-stock item procurement at IR. "Don't you have a unique identifier to locate a non-stock item?" was Reddy's query after he had understood the background of the problem. It was then that Reddy realized that he needed to cluster items based not only on their lexical similarity but also on their semantic similarity. If one can successfully group similar items in the same cluster, it will achieve two purposes. First, it will help identify past procurements of similar non-stock items. Second, it will also help field officers find stock items with identical text descriptions that had been inadvertently assigned different PL numbers.

AI promised to open tremendous opportunities for IR. From improving patient safety to remote monitoring to automating cruise control, AI could impact railway operations significantly. Reddy had previously worked on consulting assignments related to AI techniques that helped firms identify relevant data points autonomously without human effort. He was confident that creating a unique identifier for each search string and clustering IMS data would help create an effective decision support system for assessing the availability and reasonableness of rates for non-stock items. Reddy also realized that character-by-character matching of search strings would not solve the problem; a more rigorous AI approach was needed that could integrate lexical and semantic similarity measures.

From his experience, Reddy realized that in natural language processing, semantic similarity techniques helped determine the similarity between texts by using their meaning rather than by character-by-character matching.[3] Many search engines used AI to identify relevant results using a "Bag of Words" (BoW)[4] and/or "Random Forests"[5] approaches rather using a keyword search.

Reddy gathered baseline data on 197 item descriptions from the materials management team of IR and used several clustering algorithms in many iterations. The predictions of different clustering algorithms were compared against the actual clusters to measure their performance accuracy (see **Exhibit 1).**

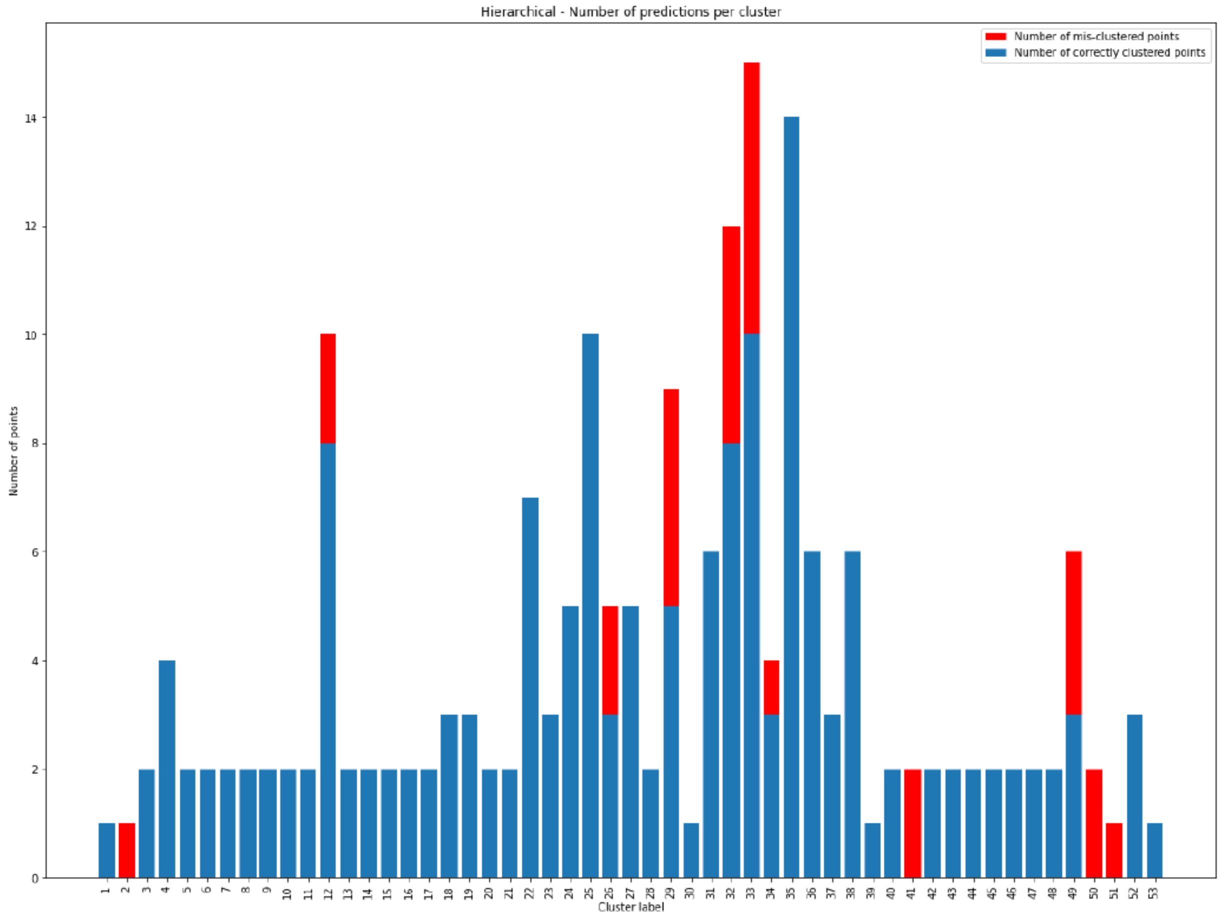
Reddy presented the preliminary results to Sumana, which showed an about 85% accuracy in matching search strings with the IMS database. He was confident that a larger data set would improve the performance of the clustering algorithm. Although Sumana was happy to see the initial results, she wondered if Reddy could solve the problem if more observations were used. As a next step, Reddy was presented baseline data with 4,000 data points (see **Exhibit 2** in the MS Excel supplement) comprising the item code, name, purchasing division, location, purchase order number, product license number, description of the non-stock item, consignee details, and time of purchase. He was tasked to get the decision support system ready for IR field workers by showing the most relevant item—based on the search string—as the first item and subsequent items with a ranking based on the similarity score.

[3] Sahami, M., & Heilman, T. D. (2006). A web-based kernel function for measuring the similarity of short text snippets. *Proceedings of the 15th International Conference on World Wide Web*, 377–386.
[4] Bag of Words (BoW) is a representation used in natural language processing in which a text is used as the bag of its words disregarding the grammar or sequence of words, but preserving the count.
[5] Random forest is an ensemble method in data mining in which a multitude of decision trees are constructed during the model training for the purpose of classification, regression, or other tasks.

Would Reddy be able to solve the problem of retrieving non-stock items from their text description? Is semantic similarity matching a good alternative to keyword-based search? What were Reddy's concerns as he set out to complete these tasks? Could the proposed algorithm solve the problem in a timely and systematic manner? These were some of the questions reverberating in the corridors of IR as Reddy was busy refining the code of his algorithm.

Exhibit 1

Sample Hierarchical Clustering Algorithm Performance



Note: The total height of each bar indicates the total number of items in the true cluster, the size of the blue bar shows the number of items in the true cluster that the model correctly predicts, and the size of the red bar indicates the number of items in the true cluster that the model misclassified. For example, true cluster 12 has ten items (total height), of which eight items were classified correctly as belonging to the same cluster (blue bar), and two items were classified as belonging to a cluster other than cluster 12 (red bar).
Source: Created by the authors.

For Exhibit 2 (Baseline Data), please refer to ISB367, the spreadsheet supplement to this case.