RESEARCH ARTICLE OPEN ACCESS

Multilingual Voice-Enabled AI Chatbots for Industrial Troubleshooting Using Unstructured Machine Manuals

Dnyanesh Badave*, Shivani Budhkar**, Abhishek Sharma***, Shubham Gugale****, Gaurav Patil*****

*(MCA Department, PES Modern College of Engineering, Pune, India
Email: dnyaneshsb20@gmail.com)
** (MCA Department, PES Modern College of Engineering, Pune, India
Email: shivanibudhkar@gmail.com)
***(TATA Motors Digital.AI Labs, Pune, India
Email: royalabhisheksharma@gmail.com)
****(TATA Motors Digital.AI Labs, Pune, India
Email: s.k.gugale@gmail.com)
*****(TATA Motors Digital.AI Labs, Pune, India
Email: gaurav.patil.ds1@gmail.com)

--************************----------------------------------

Abstract:

Smart manufacturing environments have witnessed rapid growth in automation and condition monitoring, yet real-time troubleshooting guidance for factory-floor personnel continues to lag behind these advances. Technical documentation for industrial equipment is predominantly available in English and employs domain-specific vocabulary that proves difficult for operators to navigate swiftly under fault conditions.

This work presents a multilingual, voice-capable AI assistant that supports equipment troubleshooting by extracting knowledge from unstructured machine documentation. The system integrates Natural Language Processing with dense vector retrieval to surface contextually relevant passages from large technical corpora. Both spoken and typed operator queries are handled, semantically compared against indexed document chunks, and translated into actionable repair guidance.

The assistant lowers the barrier to technical knowledge and helps maintenance personnel resolve equipment faults with greater speed and confidence. Key engineering challenges addressed include cross-lingual communication in noisy shop-floor conditions and knowledge extraction from poorly structured documentation. Evaluation outcomes suggest that dialogue-based AI can meaningfully reduce reliance on specialist technicians, shorten equipment downtime, and raise overall plant productivity.

The study underscores the importance of human-centered intelligent systems that combine voice interaction, cross-lingual support, and semantic document retrieval within Industry 4.0 manufacturing workflows.

Keywords — Multilingual Chatbot, Industrial Troubleshooting, NLP, RAG, Smart Manufacturing

--************************----------------------------------

I. Introduction

Unplanned equipment stoppages cost manufacturers an estimated $50 billion annually in lost production, yet the tools available to shop-floor operators for diagnosing faults have changed little over the past decade[8]. A technician standing before a halted machine must still leaf through multi-hundred-page manuals written in technical English — a language many operators do not speak fluently — under the pressure of every passing minute of downtime. The result is a persistent and costly gap between the knowledge locked inside documentation and the operator who urgently needs it.

Even minor component failures carry the risk of cascading disruptions to production schedules and measurable losses in output. A critical gap in current manufacturing environments is the lack of interactive, context-sensitive tools that can deliver targeted technical guidance directly to operators at the point of need. While documentation exists in abundance, its practical utility during time-critical troubleshooting scenarios remains largely unrealised.

Progress in large language models and retrieval-based AI has made it feasible to construct conversational agents that interpret operator queries phrased in everyday natural language and return precise, document-grounded answers[8]. Such agents can search across thousands of pages of technical content and synthesise focused responses within seconds. However, the majority of deployed solutions target well-structured corpora or general-purpose dialogue, whereas industrial fault resolution from unformatted equipment manuals remains comparatively unexplored.

A further complication arises from the linguistic diversity of shop-floor workforces, where operators routinely prefer to communicate in their native regional languages rather than English. Embedding multilingual and speech-based interaction within AI retrieval pipelines therefore has significant potential to increase system adoption and practical utility for frontline maintenance personnel.

This paper proposes a multilingual AI-powered assistant that enables equipment fault resolution through semantically-driven search over unstructured technical manuals. The system employs dense vector representations and a Retrieval-Augmented Generation architecture to locate pertinent document passages and formulate precise, context-sensitive answers[8, 1]. The primary goals are to shorten mean time to repair, reduce dependency on scarce expert knowledge, and make technical documentation genuinely accessible to all operators regardless of language background.

II. Literature Review

The maturation of large language models and neural retrieval architectures has opened new avenues for building knowledge-intensive conversational systems capable of operating over large collections of unformatted text. Generative models augmented with external retrieval mechanisms have demonstrated substantial gains in factual grounding and a marked reduction in confabulated responses compared with purely parametric approaches[3, 7, 8].

A variety of architectural strategies have been investigated to strengthen chatbot reliability and answer quality. Hybrid pipelines that combine dense semantic search with a bank of pre-generated question-answer pairs consistently outperform retrieval-only or generation-only baselines on both precision and user satisfaction metrics[1]. Complementary to this, score-weighted retrieval schemes allow systems to privilege the most contextually aligned passages, a property especially valuable in technical support scenarios where precision is critical[2].

The RAG paradigm has been successfully deployed across a range of application domains, including academic tutoring platforms, enterprise knowledge management, and web-scale information services[3, 4, 5]. Each of these deployments validates the utility of vector-indexed semantic search as a mechanism for unlocking value from previously inaccessible unstructured knowledge stores. Attention has also been directed toward principled evaluation frameworks that measure retrieval fidelity, response coherence, and end-to-end system reliability in deployment conditions[6, 7].

Systematic comparisons of retrieval strategies and backbone language models have yielded practical guidance on architecture selection and hyperparameter optimisation for dialogue systems[7]. Nonetheless, the overwhelming majority of existing systems assume English-language input and rely on well-formatted source documents. Multilingual conversational interfaces grounded in raw industrial documentation represent an underexplored area, and the convergence of voice interaction, cross-lingual retrieval, and unstructured manual parsing for manufacturing troubleshooting constitutes a clear and pressing research opportunity.

III. Research Gap and Challenges

Although RAG-based conversational AI has matured rapidly, its penetration into industrial maintenance and fault diagnosis workflows remains shallow[8]. Existing deployments are predominantly oriented toward consumer-facing support services, academic question-answering, or structured enterprise databases rather than the unformatted, domain-specific documentation characteristic of manufacturing environments[3, 4, 5].

Equipment technical manuals typically comprise heterogeneous, loosely organised content spanning installation procedures, operating limits, calibration steps, and fault code tables. Conventional keyword-based retrieval architectures struggle with this format and frequently return incomplete or misleading results[1, 2]. Additionally, most commercially available chatbot platforms offer limited or no support for regional language input, which directly excludes a significant proportion of the global manufacturing workforce from the benefits of AI-assisted troubleshooting.

Voice-based interaction presents a separate set of open research challenges, particularly in acoustically harsh shop-floor environments where background machinery noise can severely degrade speech recognition accuracy. Sustaining high-quality, equipment-specific knowledge bases and ensuring that retrieval results remain contextually aligned with the operational state of a specific machine are additional unsolved problems.

A fundamental shortcoming of many deployed retrieval systems is their reliance on lexical overlap signals rather than a genuine understanding of operational context and user intent[2]. Addressing this limitation requires a new generation of industrial AI assistants that integrate multilingual voice input, context-aware semantic retrieval, and adaptive response generation as first-class design requirements.

IV. Methodology

A. Data Collection

The experimental corpus comprises maintenance handbooks, fault diagnosis guides, and operational specifications for a range of industrial machinery, predominantly sourced as PDF documents.

Text was extracted from each document using automated parsing tools and subjected to noise removal and normalisation. The cleaned text was then partitioned into overlapping semantic windows of fixed token length, and each window was encoded as a dense embedding vector using a pretrained sentence encoder. The resulting vectors were inserted into a FAISS index[9] along with positional metadata to support sub-second nearest-neighbour retrieval at query time.

B. Proposed Methodology

The proposed system operates as a multilingual industrial troubleshooting assistant by coordinating NLP preprocessing, dense vector retrieval, and large language model generation into a unified pipeline.

On receiving an operator query, the system first applies language identification and, if necessary, translates the input into English using a neural machine translation module. The translated text is then encoded into the same vector space as the document index to enable semantic similarity comparison. A ranked list of the most contextually relevant document passages is retrieved from the FAISS index[9].

These passages are concatenated with the original query and supplied as context to a large language model, which synthesises a focused, jargon-aware troubleshooting response following the RAG paradigm[8]. Where the operator's preferred language is not English, the generated response undergoes a final back-translation step before display.

This architecture ensures that relevant technical content is surfaced rapidly from unstructured documents and communicated to operators in a form that is both linguistically accessible and operationally actionable.

V. Implementation and Design

The working prototype is a Python-based troubleshooting platform that coordinates document ingestion, semantic indexing, and generative response synthesis into a cohesive end-to-end pipeline. PDF documents are parsed using text extraction libraries; the resulting content undergoes cleaning and segmentation before being encoded as dense vectors and stored in a FAISS index[9] for efficient similarity-based lookup.

Multiple chunking granularities and overlap strategies were benchmarked during development to identify the configuration that maximised retrieval precision across a diverse query set. A plug-in translation component, based on a pretrained multilingual model, allows operators to pose queries and receive responses in a range of regional languages without any changes to the underlying retrieval or generation modules.

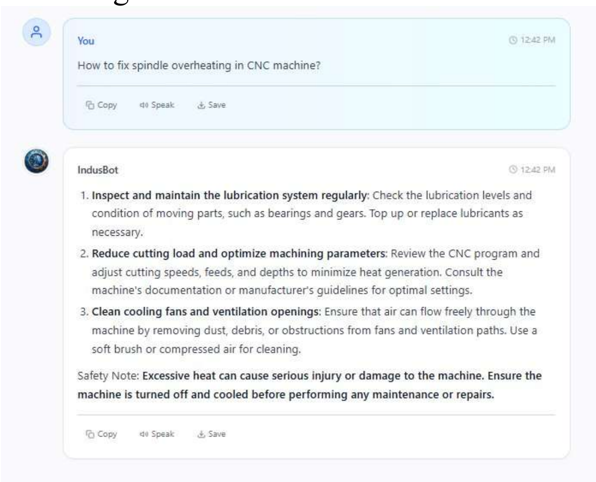

Fig. 1 IndusBot Response for CNC Spindle Overheating Issue

At inference time, a submitted query is vectorised and compared against the indexed embeddings to retrieve the top-*k* document passages. These passages are assembled into a structured prompt and passed to the language model[8], which produces a concise, situation-specific response. The answer is rendered in the chat interface, allowing operators to obtain precise fault resolution guidance without manually consulting multi-hundred-page manuals.

The prototype confirms that pairing semantic retrieval with generative language modelling yields a measurable improvement in the accessibility and relevance of technical knowledge for industrial users.

VI. Results and Discussion

Empirical assessment of the system involved submitting a battery of equipment-specific fault queries spanning multiple machine types and failure modes. The results indicate that grounding generative responses in semantically retrieved document passages substantially improves both the precision and operational applicability of troubleshooting guidance[8].

The FAISS-based retrieval layer[9] demonstrated reliable identification of pertinent manual sections even for queries that shared no lexical overlap with the source text, confirming the advantage of embedding-based similarity over conventional inverted-index search[1, 2].

Cross-lingual query handling proved effective: regional-language inputs were correctly identified, translated, and matched against the English-language index, with the final response accurately back-translated for the operator. This capability closes a meaningful accessibility gap in multilingual manufacturing environments.

The conversational interface further reduced cognitive load on operators by distilling lengthy procedural documentation into concise step-by-step guidance. Generative summarisation by the large language model[3, 7] allowed responses to be tailored to the specific fault context rather than returning raw manual excerpts.

Retrieval quality was observed to degrade for ambiguous or semantically sparse queries and for source documents with inconsistent structure or missing section headings, pointing to the importance of document preprocessing quality[6]. These observations inform the improvements planned for the next development iteration.

VII. Conclusion and Future Scope

C. *Conclusion*

This paper has presented a multilingual AI assistant engineered to support real-time fault resolution in manufacturing settings that rely on unstructured technical documentation. The system integrates neural language understanding, dense vector indexing, and large language model generation to extract, rank, and communicate relevant maintenance knowledge from raw equipment manuals.

Grounding the assistant in the RAG framework[8] transforms static, difficult-to-navigate documentation into a conversational resource that delivers targeted guidance in response to each operator's specific query. Support for multiple languages extends this capability to workforce populations that have historically been excluded by English-only documentation.

The experimental findings affirm that conversational AI architectures are a viable and impactful mechanism for improving technical knowledge access, reducing unplanned downtime, and strengthening maintenance operations across Industry 4.0 manufacturing environments.

D. *Future Scope*

Planned extensions to the current system include native voice input with noise-robust speech recognition calibrated for factory acoustic conditions, which will further eliminate barriers to adoption on the shop floor. Streaming integration with IoT sensor feeds from connected equipment offers the prospect of real-time, condition-aware diagnostic suggestions that go beyond static manual content[9].

Further research directions encompass expanding language coverage to additional regional languages, investigating advanced chunking and re-ranking strategies to improve retrieval recall, and coupling the assistant with enterprise asset management and ERP platforms to create a fully integrated maintenance intelligence layer. These advances would consolidate conversational AI as a foundational capability within smart manufacturing ecosystems.

References

[1] S. Kim, H. Jeon, D. Kim, M. Kim, D.-K. Chae, and J. Kim, "HybridRAG: A Practical LLM-based ChatBot Framework based on Pre-Generated Q&A over Raw Unstructured Documents," *arXiv preprint arXiv:2602.11156*, 2025.

[2] R. Khanda, "Agentic AI-Driven Technical Troubleshooting for Enterprise Systems: A Novel Weighted RAG Paradigm," *arXiv preprint arXiv:2412.12006*, 2024.

[3] J. Swacha and M. Gracel, "Retrieval-Augmented Generation (RAG) Chatbots for Education: A Survey of Applications," *Applied Sciences*, vol. 15, no. 8, pp. 4234, 2025.

[4] V. Yadav, A. Yadav, S. Yadav, A. Birje, and M. Gupta, "RAG Based Personal AI Chatbot for College Information System," *International Journal for Research in Applied Science & Engineering Technology (IJRASET)*, vol. 13, no. 5, 2025.

[5] B. Sarmah, B. Hall, R. Rao, S. Patel, S. Pasquali, and D. Mehta, "A Practical Application of RAG for Website-Based Chatbots Combining Scraping, Vectorization & Semantic Search," *Journal of Theoretical and Computational Science and Software Technology*, vol. 4, no. 4, 2024.

[6] M. T. Islam, S. S. Islam, and J. Shin, "Retrieval Augmented Generation (RAG) Based Restaurant Chatbot with AI Testability," in *Proc. IEEE International Conference on Big Data Service and Applications*, 2024.

[7] A. Alzahrani and K. Salama, "Retrieval-Augmented Generation (RAG) Chatbots: A Comparative Study," San Jose State University ScholarWorks Repository, 2024.

[8] P. Lewis, E. Perez, A. Piktus, F. Petroni, V. Karpukhin, N. Goyal, H. Küttler, M. Lewis, W. Yih, T. Rocktäschel, S. Riedel, and D. Kiela, "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks," in *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 33, pp. 9459–9474, 2020.

[9] J. Johnson, M. Douze, and H. Jégou, "Billion-Scale Similarity Search with GPUs," *IEEE Transactions on Big Data*, vol. 7, no. 3, pp. 535–547, 2019.

