PROBLEM STATEMENT

TABLE DOCUMENT - 1

AUTOMATED TITLE VERIFICATION AND REGISTRATION SYSTEM FOR PRGI

PROBLEM STATEMENT TITLE	An online system to automatically verify new title submissions by checking for similarities with existing titles.
CONTEXT	Press Registrar General of India (PRGI) maintains a database containing approximately 160,000 titles. When a user submits a new title for verification, we need to check its similarity against the existing titles in our database. The goal is to ensure that the new title does not duplicate or closely resemble any existing title to avoid confusion and maintain uniqueness. Additionally, the system must enforce specific guidelines to ensure that certain words are not used, combinations of existing titles are not allowed, and titles with similar meanings or periodicity modifications are rejected. However, manual verification is time-consuming, and variations in spelling, phonetics, and meaning make similarity detection challenging. A systematic approach is required to automate the verification process while enforcing rules against restricted words, title combinations, and periodicity modifications.
ROOT CAUSE	The primary root cause of the problem is the absence of an automated and intelligent title verification system that can efficiently handle large-scale comparisons. The existing database lacks an advanced similarity-checking mechanism, making manual verification inefficient, slow, and prone to human errors. Additionally, the lack of predefined rule enforcement within the system leads to inconsistencies in title approval. The challenge is further compounded by variations in language, synonyms, and minor modifications that make it difficult to detect near-duplicates accurately.
IDEAL SOLUTION	The ideal solution is to develop an AI-powered automated title verification system that integrates natural language processing (NLP) and machine learning algorithms to efficiently compare new submissions with the existing database. This system

	should include fuzzy matching techniques to detect near-duplicates, rule-based filters to enforce predefined guidelines, and context-aware analysis to identify semantic similarities. Additionally, the solution should incorporate a real-time feedback mechanism to provide immediate validation results to users, reducing the need for manual intervention. A user-friendly interface and continuous model training will further enhance accuracy and efficiency over time.
PROPOSED SOLUTION	To address the challenge of verifying new title submissions, we propose developing an automated system that checks for similarities with existing titles in the database while ensuring compliance with predefined guidelines. The system will employ phonetic similarity algorithms (e.g., Soundex, Metaphone) to detect similar-sounding names, identify common prefixes or suffixes, and prevent slight spelling variations from bypassing verification. It will enforce restrictions on disallowed words, prohibit the combination of existing titles, and reject titles with similar meanings in other languages. Additionally, a probability score will be generated to indicate the likelihood of a title being verified. The system will leverage efficient search and indexing techniques to handle large-scale data, ensuring quick responses and scalability. Users will receive clear feedback on rejections, including explanations for similarity violations, disallowed terms, or guideline breaches, and will be allowed to modify and resubmit titles for verification. Furthermore, the system will track current applications to prevent duplicate submissions over time. Designed for accuracy, performance, and user experience, the solution aims to deliver rapid, reliable title verification (and registration) while maintaining database integrity.

END