Research on Technology Opportunity Identification Based on Technical Boundaries

Heng Chen, Wenbo Zhong*

School of Economics and Management, Harbin Engineering University, Harbin, China

Email address:

chenh@hrbeu.edu.cn (Heng Chen), s944307951@163.com (Wenbo Zhong)

*Corresponding author

Abstract

Identifying authentic technology opportunities from massive patent data is a critical task for strategic R&D and competitive intelligence. Traditional approaches predominantly rely on structured classification systems such as IPC or CPC, which often struggle to capture the granular, evolving, and cross-disciplinary nature of modern technical intelligence. This research presents a systematic framework for identifying technology opportunities by constructing and analyzing mathematically defined "Technical Boundaries" from unstructured textual data. The methodology begins by leveraging BERTopic combined with the BGE-M3 embedding model to transform patent corpora into a high-dimensional semantic manifold, preserving intricate technical features. To address the inherent challenge of data sparsity in high-dimensional spaces and ensure semantic consistency, a Parametric UMAP architecture is implemented. Unlike traditional non-parametric dimension reduction, this deep-learning-based approach utilizes an Encoder-Decoder mechanism to establish a stable, bi-directional mapping between low-dimensional latent coordinates and high-dimensional technical features. Central to the framework is the application of the One-Class SVM (OC-SVM) algorithm to demarcate a "Normative Technical Boundary" based on established CPC-defined knowledge. By quantifying the density distribution and the spillover effects relative to this boundary, the research categorizes technology opportunities into two distinct types: "Filling-type Opportunities", representing critical technical voids within existing frameworks, and "Expansionary-type Opportunities", signifying pioneering breakthroughs beyond established frontiers. Finally, an Inverse Mapping (reconstruction) mechanism is employed, using the Decoder to translate identified spatial voids back into interpretable technical insights. Empirical results demonstrate that this boundary-constrained methodology effectively filters semantic noise and provides a more rigorous, generative path for technological forecasting, offering strategic guidance for navigating complex innovation landscapes.

Keywords

Technology Opportunity, Technical Boundary, Parametric UMAP, BERTopic, One-Class SVM, Patent Mining