

INNOVATION FOR IMPACT: COMPREHENDING THE SUSTAINABILITY STARTUP LANDSCAPE IN INDIA

ABSTRACT

Over the past decade, the Indian startup ecosystem has undergone a significant transformation, emerging as one of the most dynamic entrepreneurial landscapes globally. This evolution has been driven by digital transformation, supportive government initiatives, expanding venture capital participation, institutional enablers such as incubators and entrepreneurial universities, and a growing pool of skilled entrepreneurial talent. Among these developments, sustainability-oriented startups have gained increasing prominence as key actors addressing pressing social and environmental challenges that align with the Sustainable Development Goals (SDGs). Positioned at the intersection of innovation, impact, and growth, these ventures are particularly salient in the Indian context, characterised by resource constraints alongside a rapidly maturing innovation ecosystem.

This thesis undertakes a comprehensive and multi-method examination of the evolution, challenges, funding dynamics, and growth trajectories of sustainability startups in India. The study begins by mapping the temporal and spatial progression of sustainability startups, revealing patterns of rapid growth, evolving funding behaviour, and the geographic concentration of SDG-focused ventures across major urban centres. The research identifies co-occurrence patterns among SDG focus areas, highlighting how clusters of sustainability innovation emerge in response to localised challenges and institutional contexts. These findings underscore the importance of context-sensitive strategies in fostering impactful sustainability entrepreneurship.

To address limitations in existing classification approaches, the research develops an automated, scalable system for SDG alignment using multi-modal data. Leveraging state-of-the-art Natural Language Processing techniques and transformer-based models such as BERT, the system performs robust multi-label SDG classification across diverse unstructured data sources, including websites, pitch decks, and multimedia content. This machine learning-driven pipeline significantly enhances the efficiency, accuracy, and consistency of SDG classification, reducing reliance on manual methods and enabling large-scale ecosystem analysis for researchers, investors, and policymakers.

The thesis further investigates the entrepreneurial and institutional dynamics shaping startup growth. Employing a mixed-methods approach that integrates Partial Least Squares Structural Equation Modelling (PLS-SEM), LightGBM, and multi-criteria decision-making tools such as TOPSIS, the study analyses data from 135 Indian sustainability startups. The results demonstrate that entrepreneurial challenges do not directly determine growth outcomes; instead, their effects are mediated by the availability and effectiveness of institutional support mechanisms, including government policies, incubators, networks, and mentorship. This highlights the central role of ecosystem support in fostering resilience and enabling startups to convert adversity into growth opportunities.

To deepen insights into funding and performance outcomes, the study analyses 735 funded Indian sustainability startups using clustering, correlation analysis, and ridge regression. The findings reveal distinct clusters of sustainability startups with varying funding and growth characteristics, and show that revenue growth is significantly influenced by employee expansion, geographic location, and funding stage. Early-stage sustainability ventures face pronounced funding constraints, while later-stage ventures benefit from increasing investor confidence and ecosystem maturity.

Recognising the labour-intensive nature of startup evaluation, particularly for sustainability metrics, the thesis introduces a Generative AI-based framework to automate the extraction of financial and impact indicators from unstructured pitch decks. Using OpenAI embeddings, FAISS, and LlamaIndex, the system efficiently retrieves key performance metrics and demonstrates substantial reductions in screening time, validated with the pitch decks of 94 Indian startups.

This research advances understanding of sustainability entrepreneurship in emerging economies by integrating theory-driven analysis with machine learning, automation and artificial intelligence. It provides a holistic, data-driven perspective on how Indian sustainability startups evolve, grow, and secure funding within complex institutional environments, offering actionable insights for entrepreneurs, investors, and policymakers committed to advancing sustainable development.

Keywords: Sustainability, Startups, SDG, Machine Learning, NLP, GenAI, Entrepreneurial Challenges, Support, Funding, Revenue Growth, Pitch Deck Evaluation