ARTPARK
AI & Robotics Technology Park, I-Hub @ IISc

KSCST
Celebrating Golden Jubilee 1975-2025

ಕರ್ನಾಟಕ ಸರ್ಕಾರ

सत्यमेव जयते

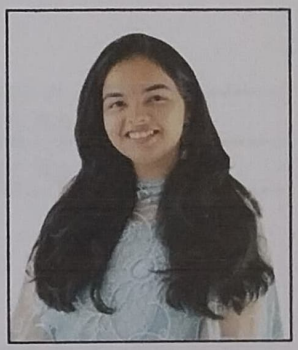
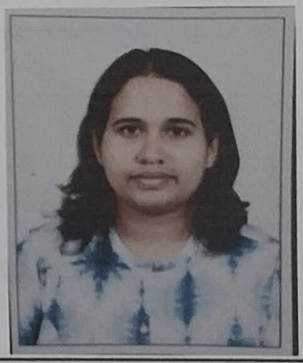
KARNATAKA STATE COUNCIL FOR SCIENCE AND TECHNOLOGY

Indian Institute of Science campus, Bengaluru

Telephone: 080 – 23341652, 23348848 || Email: spp@kscst.org.in
Website: www.kscst.org.in/spp.html or https://kscst.karnataka.gov.in/en

FORMAT FOR STUDENT PROJECT PROPOSAL FOR THE 49th SERIES OF STUDENT PROJECT PROGRAMME

1.	**Name of the College: KLS Gogte Institute Of Technology**
2.	**Project Title: Intelligent Multi-Marketplace Price Optimization Using Single-Agent and Multi-Agent Systems**
3.	**Branch: Electronic and Communication**
4.	**Theme (as per KSCST poster): Artificial Intelligence**
5.	**Name(s) of project guide(s):** 1. **Name: Dr.Praveen Kalkundri** **Email id: pukalkundri@git.edu** **Contact No.: 9035072685**
6.	**Name of Team Members :** **Name: ANURAG UGARGOL** **USN No.: 2GI22EC023** **Email id: 22u1639@students.git.edu** **Mobile No: 9141063645** 

	Name: APARNA TUDAVEKAR **USN No.: 2GI22EC024** **Email id: 22u1628@students.git.edu** **Mobile No.: 8762548987**  **Name: CHARUTA JOSHI** **USN No.: 2GI22EC039** **Email id: 22u1084@students.git.edu** **Mobile No.: 9980934672**  **Name: GOURI NAYAK** **USN No.: 2GI22EC50** **Email id: 22u1577@students.git.edu** **Mobile No.: 7411151965** 
7.	**Team Leader of the Project:** **Name: ANURAG UGARGOL** **USN No.: 2GI22EC023** **Email id: 22u1639@students.git.edu** **Mobile No.: 9141063645**
8.	**Processing Fee Details (Through Online Payment only):** **(processing fee of Rs. 1180/-)** **Please furnish the payment details in the format provided in the last page of the proposal.**
9.	**Date of commencement of the Project:05/09/2025**
10.	**Probable date of completion of the project: 20/04/2026**

11. **Timeline Structure**

Task	Start Date	End Date	Duration	Milestone achievable
Project Initiation	05/09/2025	05/10/2025	30 Days	Yes
Research Phase	05/10/2025	10/12/2025	65 Days	Yes
Development Phase	10/12/2025	20/02/2026	50 Days	No
Testing Phase	20/02//2026	06/04/2026	46 Days	No
Submission of PPT and Abstract to KSCST for conducting mid-term evaluation	Middle of March – 2nd Week of April 2026			
Final Review	06/04/2026	26/04/2026	20 Days	No
Project Submission	26/04/2026	02/05/2026	6 days	No
Submission of Project Completion Report to KSCST (Upload to Google Form)	May 2026			

12. **Scope / Objectives of the project:**

1.Multi-Platform Product Aggregation:

- Scrape and collect product data from Amazon.in and Flipkart simultaneously for unified comparison.
- Extract comprehensive details including pricing, technical specifications, features, and rating breakdowns.

2.RAG-Based Smart Caching:

- Implement a Retrieval-Augmented Generation (RAG) pipeline to store and retrieve product data locally.
- Use semantic search to reduce redundant web scraping and improve response times.

3.Machine Learning Sentiment Analysis:

- Train a Logistic Regression classifier on Amazon review datasets.
- Analyze customer sentiment (positive, neutral, negative) and generate sentiment scores for products.

4.Intelligent Data Processing:

- Apply TF-IDF vectorization and Cosine Similarity for fuzzy product matching.
- Filter accessories and irrelevant products automatically to ensure search relevance.

5.Interactive Visualization:

- Develop a desktop GUI using Tkinter to display product comparisons, images, specs, and sentiment insights.

13. **Methodology:**

1. System Architecture Setup:

- Design a modular architecture linking the Tkinter GUI, RAG storage pipeline, and Machine Learning sentiment analyzer.
- Implement a caching layer using Pickle persistence for fast data retrieval.

2. RAG Pipeline Implementation:

- **Local Search:** check the local database for exact matches first.
- **Fuzzy Matching:** perform semantic search using TF-IDF and Cosine Similarity with a 60% token overlap threshold if no exact match exists.
- **Storage:** Cache new search results automatically to the local RAG database.

3. Web Scraping & Data Extraction:

- **Browser Automation:** Deploy Selenium WebDriver with Chrome to handle dynamic content.
- **Anti-Bot Measures:** Implement random User-Agent rotation, delays, and window management to bypass bot detection.
- **Parallel Execution:** Scrape Amazon.in and Flipkart concurrently to minimize wait times.
- **Deep Extraction:** Parse multiple CSS selectors to retrieve prices, descriptions, specifications, and review counts.

4. Sentiment Analysis Model Development:

- **Preprocessing:** Clean text by handling negations (e.g., converting "not good" to "not good_NEG") and expanding contractions.
- **Feature Engineering:** Use TF-IDF vectorization with unigrams and bigrams (max features: 5000).
- **Training:** Train a Logistic Regression model with balanced class weights using the Amazon reviews dataset.
- **Scoring:** Convert model outputs into sentiment labels and numerical confidence scores.

5. GUI Deployment & Integration:

- Build a user interface to accept queries and display side-by-side product comparisons.
- Integrate image processing using Pillow (PIL) to render product.

14. **Expected Outcome of the project:**

1. **Unified Price Comparison Platform:** A single interface to compare real-time prices, specifications, and ratings across Amazon.in and Flipkart, eliminating the need to browse multiple sites manually.
2. **AI-Driven Consumer Insights:** Actionable sentiment analysis (Positive, Neutral, Negative) derived from customer reviews to help users judge product quality beyond just star ratings.
3. **Optimized Search Performance:** fast and efficient data retrieval achieved through a RAG (Retrieval-Augmented Generation) pipeline that caches results locally, significantly reducing search latency and redundant network requests.
4. **Comprehensive Product Analytics:** Detailed extraction and display of technical specifications, feature lists, and rating breakdowns for informed decision-making.
5. **Scalable Desktop Application:** A robust GUI-based tool capable of handling complex queries, filtering irrelevant products, and adapting to varying e-commerce layouts.

15. **Is the project proposed relevant to the Industry / Society or Institution?**

Yes / No: NO

16. **Can the product or process to be developed in the project be taken up for filing a Patent?**

Yes / No: NO

Prior Art search done?

Yes/No: YES

17. **Budget details (break-up details should be given):**

Budget	Amount
a) Materials / Consumables (T4-GPU Goolge Colab-For Training / API Calls)	2,500.00
b) Labour (Self- Work)	0.00
c) Travel (Visiting Charges)	3,000.00
e) Miscellaneous (Report printing, binding, and documentation)	500.00
Total	6000.00

18.	**Any other technical details (Please specify):** 1. **Programming Language:** Python 3.8+ 2. **Web Scraping:** Selenium WebDriver, Chrome Driver 3. **Machine Learning:** scikit-learn, Logistic Regression 4. **Data Processing:** Pandas, NumPy 5. **Natural Language Processing:** TF-IDF Vectorizer, Cosine Similarity 6. **GUI Framework:** Tkinter 7. **Image Processing:** Pillow (PIL) 8. **Version Control:** Git / GitHub
19.	**SPP Coordinator (Identified by the college):** **Name: Prof. / Dr.(Mrs) Shweta Goudar** **Email id: dean_rnd@git.edu** **Contact No.: 9448483351**

Name of the Project Guide:
Dr.Praveen Kalkundri
Email id: pukalkundri@git.edu
Contact No.: 9035072685

Name of the HOD:
Dr. Supriya Shanbhag
Email id.: ssshanbhag@git.edu
Contact No.: 9145306767