Team Name: heckers

heckers

Member 1 - Japneet Singh

Member 2 - Naman Soni

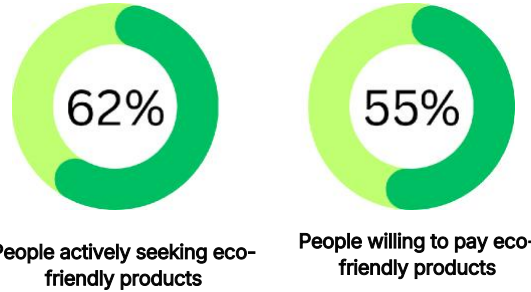
Member 3 - Sandeep Soni

Theme - Sustainable Shopping Experience

PROBLEM STATEMENT

Online shopping brings convenience but hides an **environmental cost** — eCommerce packaging cuts down **~3.2** billion trees annually and contributes to **~3%** of global carbon emissions

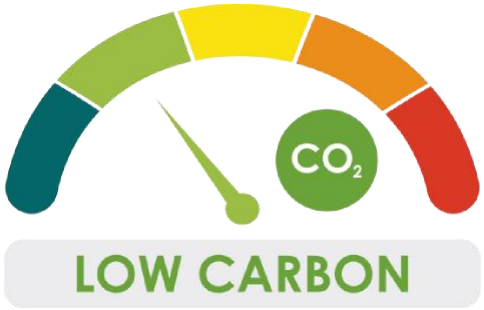
Despite **62%** of shoppers actively seeking eco-friendly products and **55–66%** willing to pay a premium, they struggle to find, compare, and trust sustainable options across platforms.



- Single-use, non-recyclable packaging and fragmented data make it hard for consumers to make green choices while shopping online.
- This growing disconnect between consumer intent and actual shopping experience poses a critical sustainability challenge in the eCommerce ecosystem.

PROBLEM STATEMENT

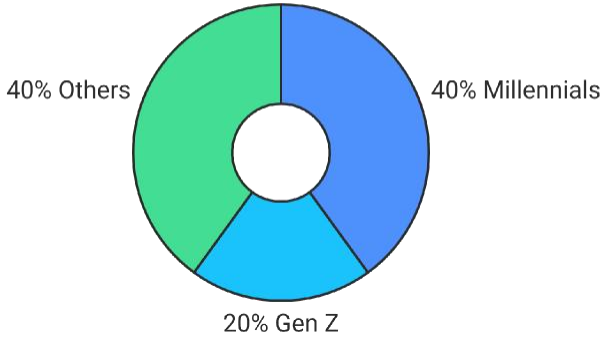
- There is a pressing need for a centralized, AI-driven system that recommends verified **eco-products**, offers **sustainable packaging choices**, and quantifies environmental impact in real time.
- Users lack access to clear **sustainability scores**, **carbon footprint data**, or aggregated impact metrics to guide conscious decision-making.
- With demand rising across groceries, personal care, home goods, and more, solving this gap can drive both environmental impact and user satisfaction.



- An AI-powered **"Green Store"** portal can revolutionize shopping by enabling transparency, impact tracking, and personalized eco-friendly experiences.
- **Innovation Scope:** AI-curated eco-products, real-time sustainability scoring, user impact dashboard, packaging and group-buying options.

Customer Profile (Working Backwards)

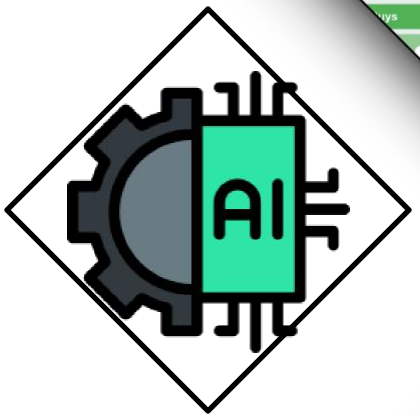
- Our primary target includes Millennials, Gen Z, and the growing segment of “sustainable mainstream” consumers — collectively **~60%** of the population
- These users are highly digital, with **46%** shopping via social media (pwc.com), and **67%** more likely to rely on online data to guide sustainable purchases
- **78%** value a sustainable lifestyle, **60%** prefer green packaging (mckinsey.com), and **66%** are willing to pay a premium for eco-products



Distribution of Eco-Conscious Individuals

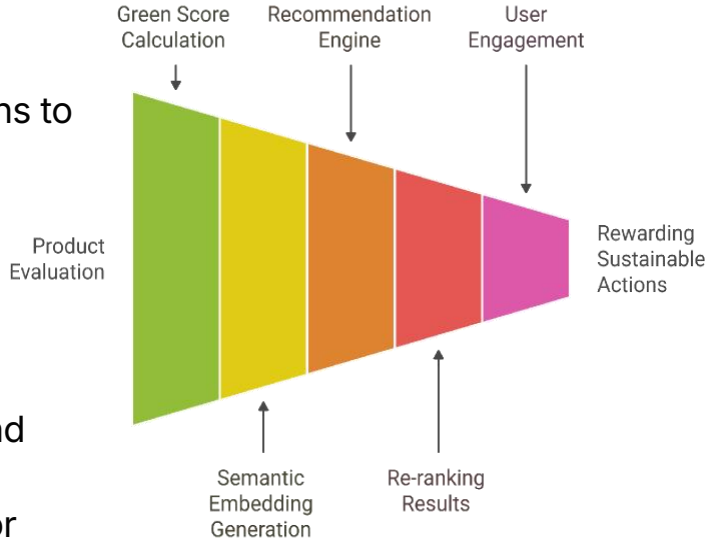
- They demand transparency — expecting visible eco-labels, verified claims (e.g. “plastic-free”), personalized carbon savings, and ethical brand values.

OUR SOLUTION



Overview – Green Score & Recommendation Engine

- Every product is evaluated using sustainability factors like carbon footprint, packaging type, and eco-certifications, leading to a *Green Score (0–100)* that reflects its environmental impact.
- A regression-based ML model, trained on expert-annotated data, computes the Green Score. At the same time, ***Sentence-BERT*** generates semantic embeddings for product titles and descriptions to enable intelligent comparisons.
- Using these embeddings, a recommendation engine suggests semantically similar products that score higher on sustainability. Results are re-ranked to prioritize greener options.
- Users can explore curated green lists such as “Top Eco Picks” and “Green Storefront,” while a personalized dashboard tracks their carbon savings and rewards them with redeemable eco-points for sustainable actions.



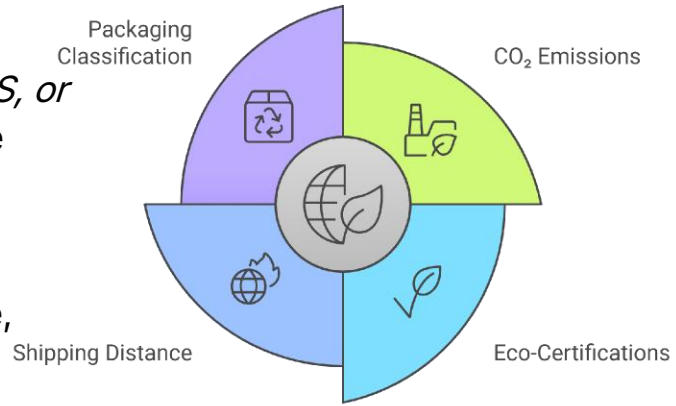
Eco Score Factors – Building Blocks of the Green Score

Core sustainability data is extracted for each product (SKU), covering lifecycle aspects from raw materials to delivery.

Estimated CO_2 emissions are calculated by analyzing inputs from manufacturing, material sourcing, and shipping logistics.

The system checks for *verified eco-certifications* such as *FSC, GOTS, or Energy Star*, which signal environmental compliance and responsible production.

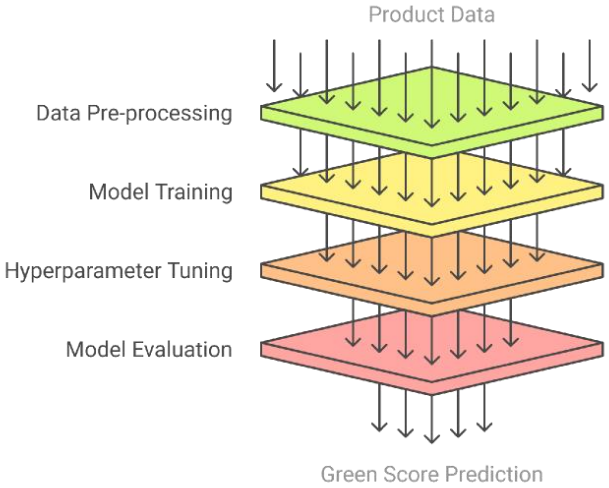
Additional features include *shipping distance estimation* (seller to customer) and *packaging classification*—whether it's fully recyclable, compostable, or mixed-material.



All attributes are selected to reflect a product's end-to-end sustainability footprint, enabling fair scoring and eco-impact analysis.

ML Model : Implementation

- A regression model *(Random Forest or XGBoost)* computes a product's *Green Score (0–100)* based on sustainability indicators.



- **Pre-processing:** Normalize numerical features like CO_2 and shipping distance using min-max scaling. Encode eco-certifications as weighted binaries and impute missing values with category-level averages.
- **Model Training:** Train on 1,000–2,000 expert-labeled products using an 80/20 train-validation split. Hyperparameters are tuned via grid search for optimal performance.
- **Evaluation & Inference:** Target performance is *Mean Absolute Error ≤ 5 and $R^2 \geq 0.8$*.
- During inference, extract 4 features per SKU, predict the *Green Score*, and store results for powering search, filters, and recommendations.

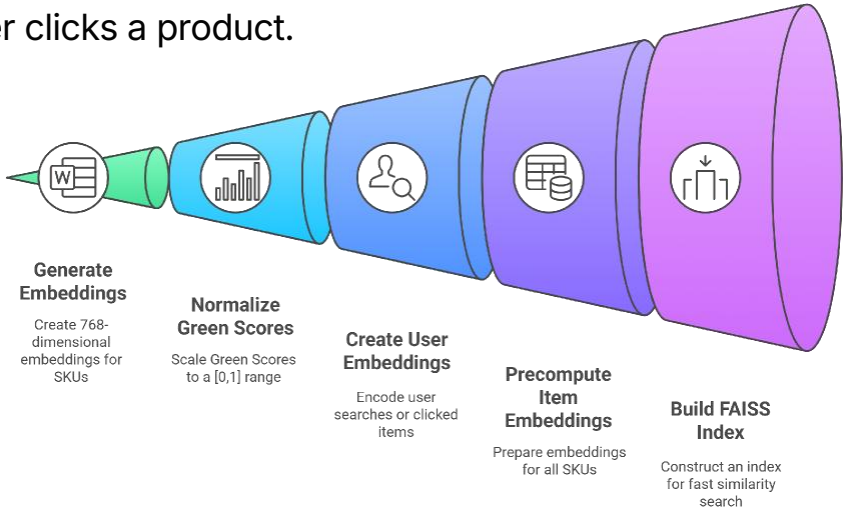
Recommendation Engine – Architecture & Inputs

1. Embedding Generation & Normalization:

- We generate *768-dimensional Sentence-BERT* embeddings for each eco-eligible SKU.
- Then normalize Green Scores by dividing by 100 to compute *green_score_norm ∈ [0,1]*.
- Create user/query embeddings by:
  - Encoding free-text searches using Sentence-BERT.
  - Using the clicked ASIN's item embedding if the user clicks a product.

2. Offline Preparation for Retrieval:

- We precompute item embeddings for all eco-eligible SKUs.
- Build a *FAISS inner-product index* over these embeddings to enable fast semantic similarity search.



Recommendation Engine – Online Workflow & Scoring

Query Processing & Retrieval:

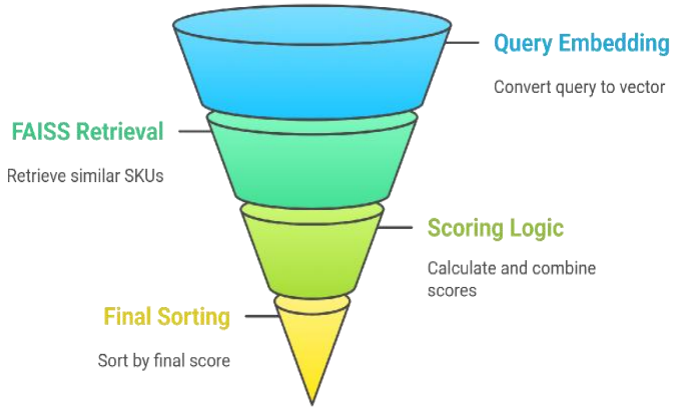
- We compute the query embedding based on user input (free-text search or clicked item).
- Using *FAISS* we retrieve the top 200 semantically similar SKUs from the indexed embeddings.

Scoring Logic:

- For each retrieved SKU, calculate:
  - *Semantic similarity* (cosine similarity between query and item embeddings)
  - Normalized Green Score (*green_score_norm ∈ [0,1]*)
- Combine the two using:

  FinalScore = α × similarity + (1 – α) × green_score_norm

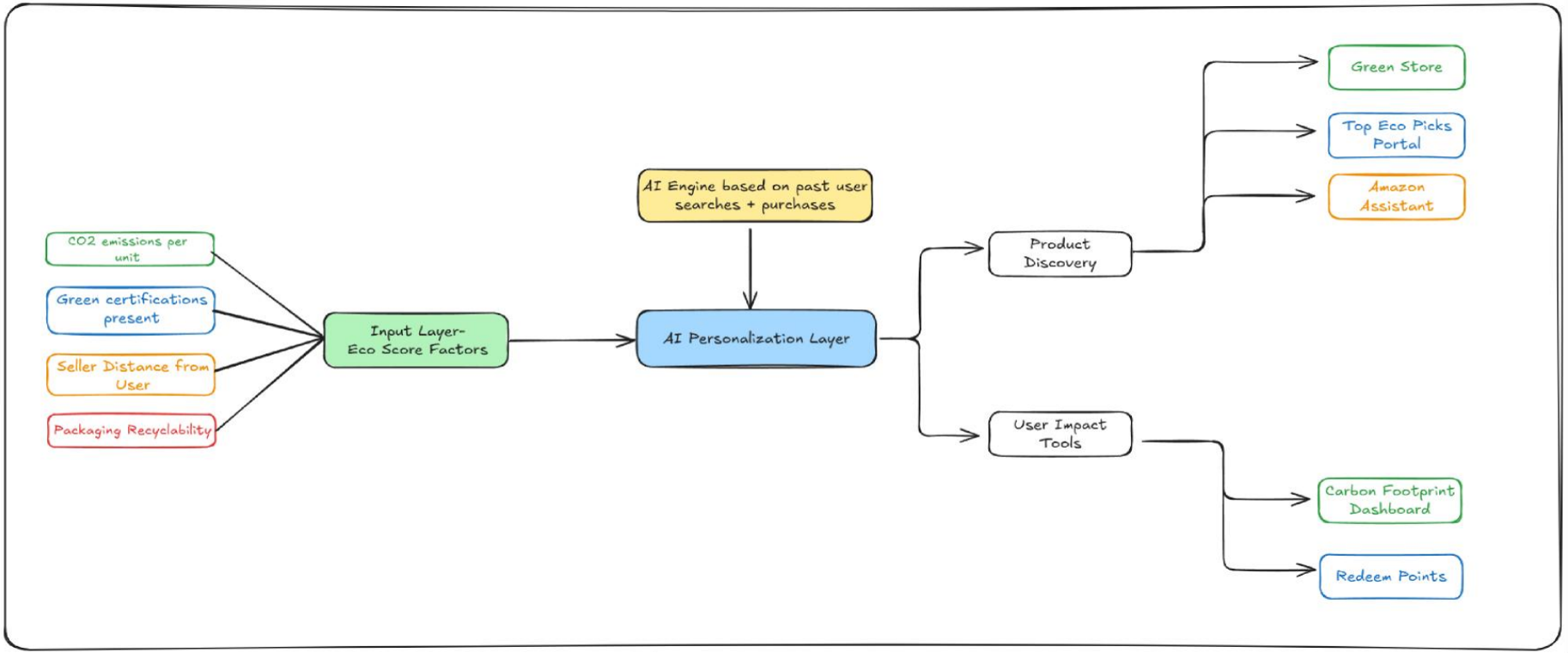
  , where α = 0.7



Recommendation Output:

- Sort the 200 candidates by Final Score (descending)
- Then return the top K results (e.g., 10) with sustainability indicators.

Flow Diagram



Recommendation Engine – Use Case & User Flow

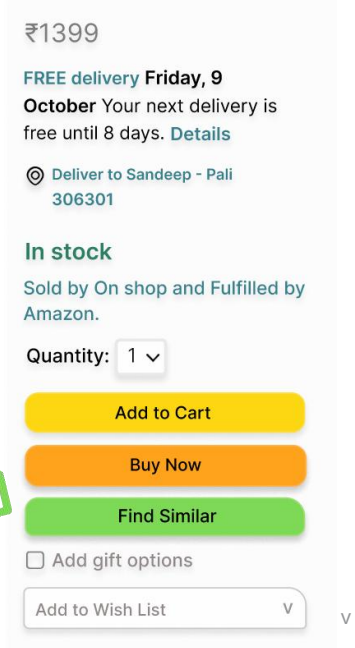
Use Case 1 – Free-Text Search:

- Users type queries like “eco-friendly water bottle.”
- The system computes a query embedding by encoding the input text.
- *FAISS* retrieves ~200 nearest SKUs based on embedding similarity.
- Candidates are reranked using: *0.7 × semantic similarity + 0.3 × green_score_norm*.
- The top 10 eco-conscious matches are shown to the user, each with a visible “Eco Score: X/100” badge

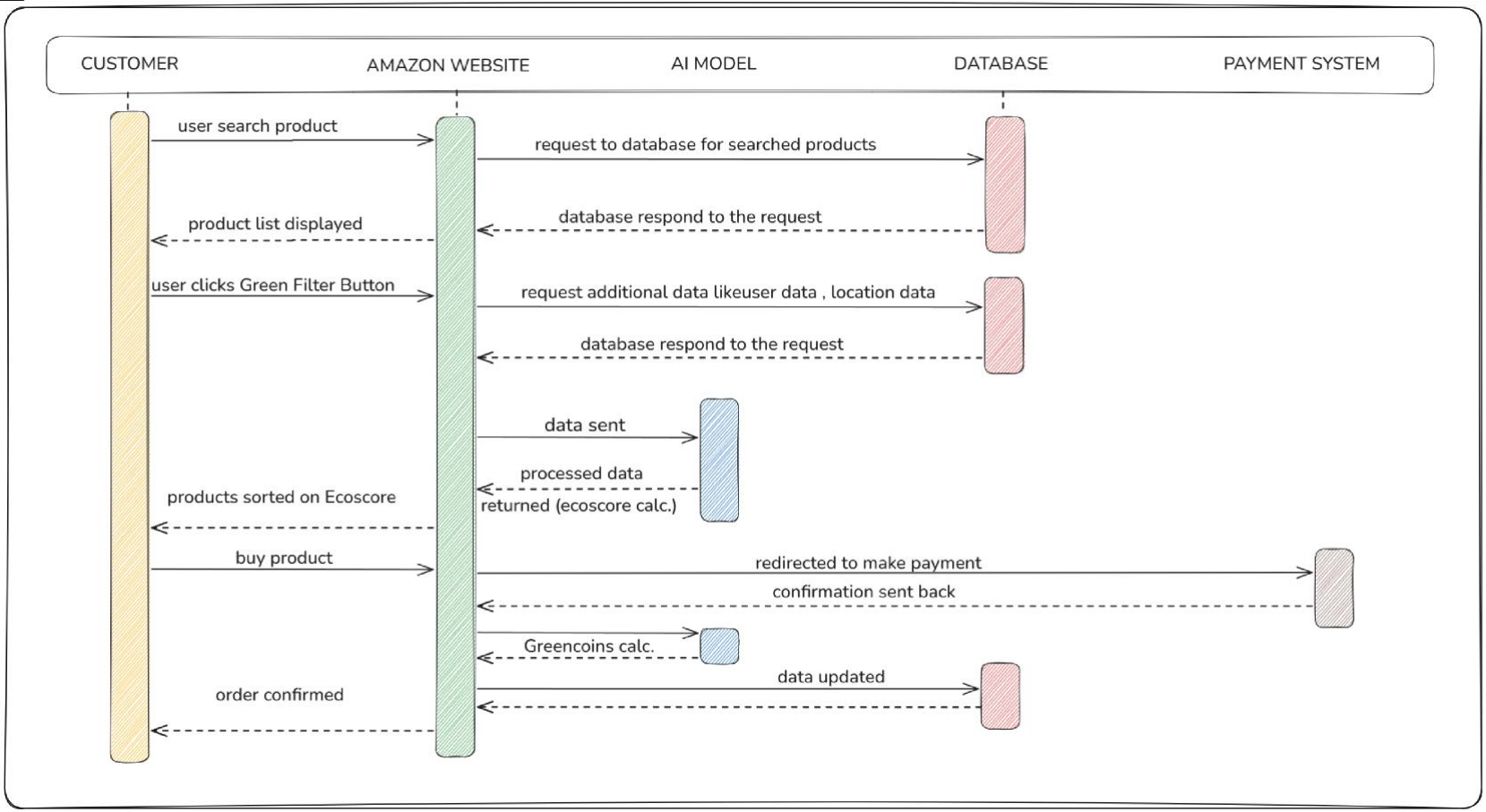
Use Case 2 – Clicked a Product:

When a user clicks “Find similar,” the system uses the product’s embedding to retrieve similar *SKUs via FAISS*.

- Results are reranked to prioritize eco-friendly alternatives, helping users switch to greener options effortlessly.



Sequence Diagram



Why These Models?

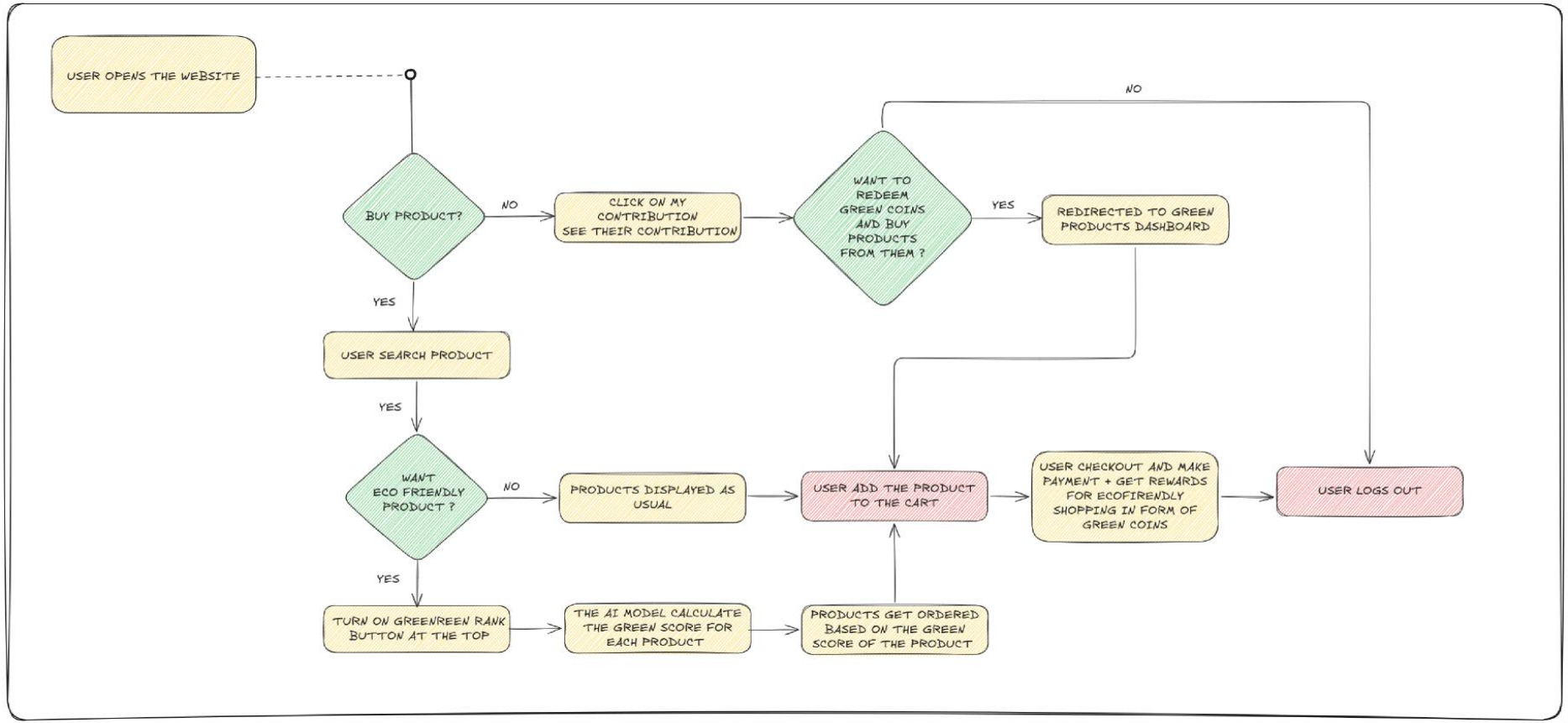
1. Green Score Model

- **Purpose:**
  Provides a transparent, quantitative assessment of product sustainability.
- **Key Benefits:**
  1. Ensures consistent eco-tagging across millions of SKUs.
  2. Enables filtering and sorting features in the “Green Store” user interfaces.
  3. Supports carbon-saving metrics on user dashboards.

1. Recommendation Engine

- **Purpose:**
  Closes the gap in traditional ranking systems that overlook environmental credentials.
- **Key Benefits:**
  1. Surfaces eco-friendly products without compromising on relevance.
  2. Highlights “top green picks” tailored to user interests, promoting informed, sustainable choices.

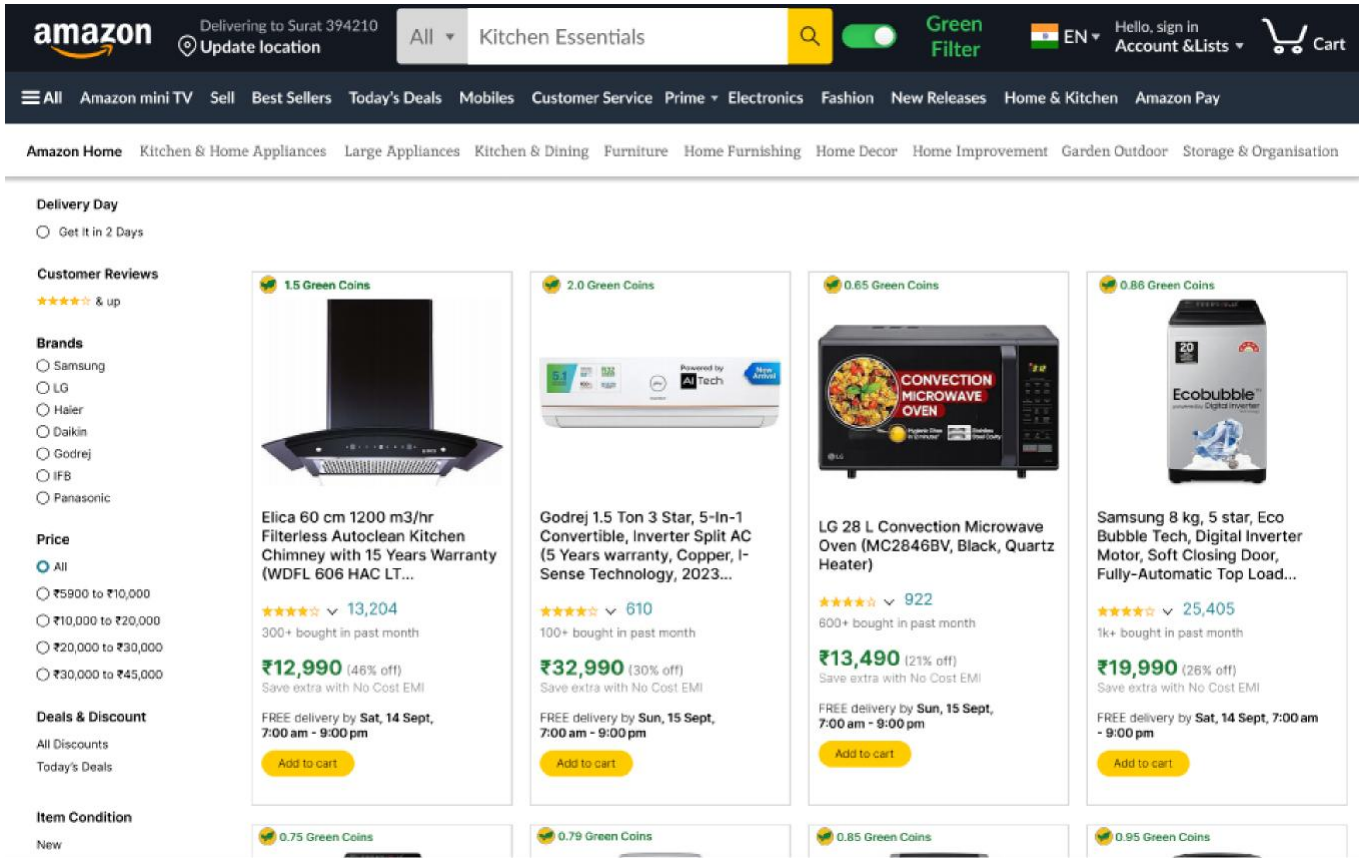
Activity Diagram



Use of Green Filter to list Eco Friendly Products

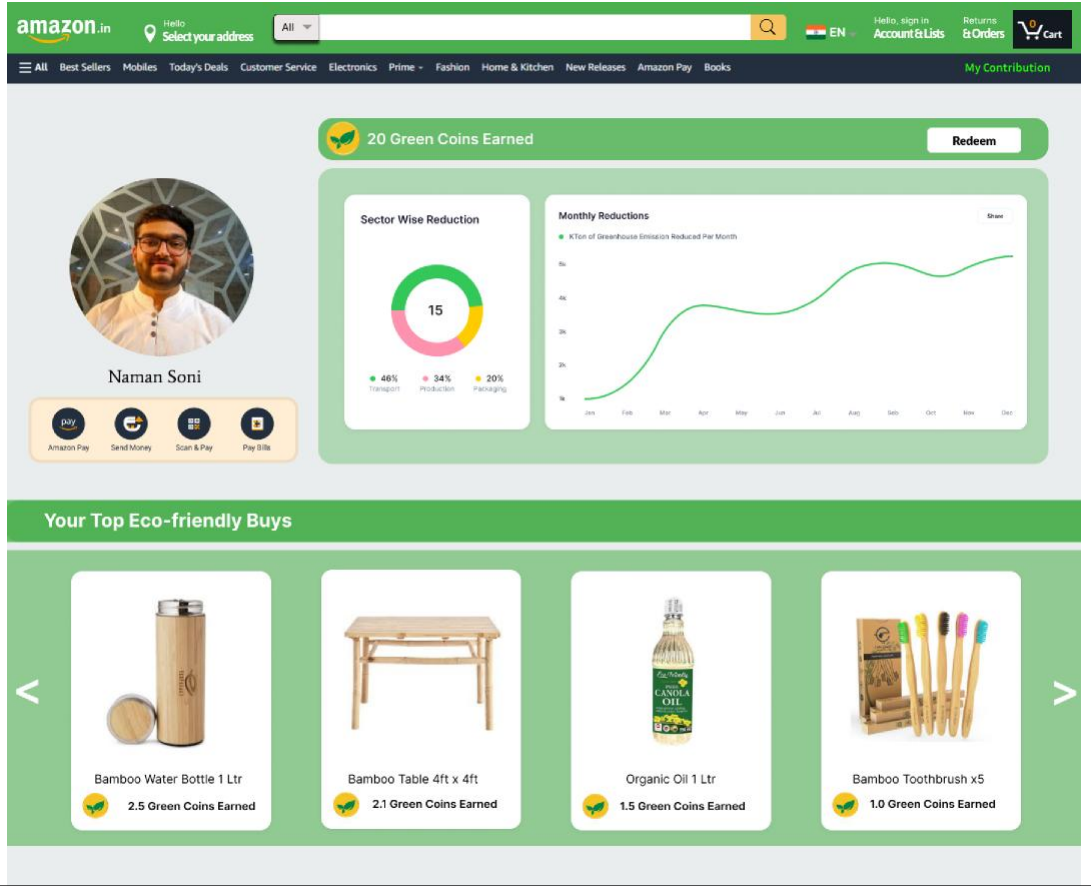
The products get displayed in sorted order based on the Green Score generated be Green Genie.

Green Filter Turned On

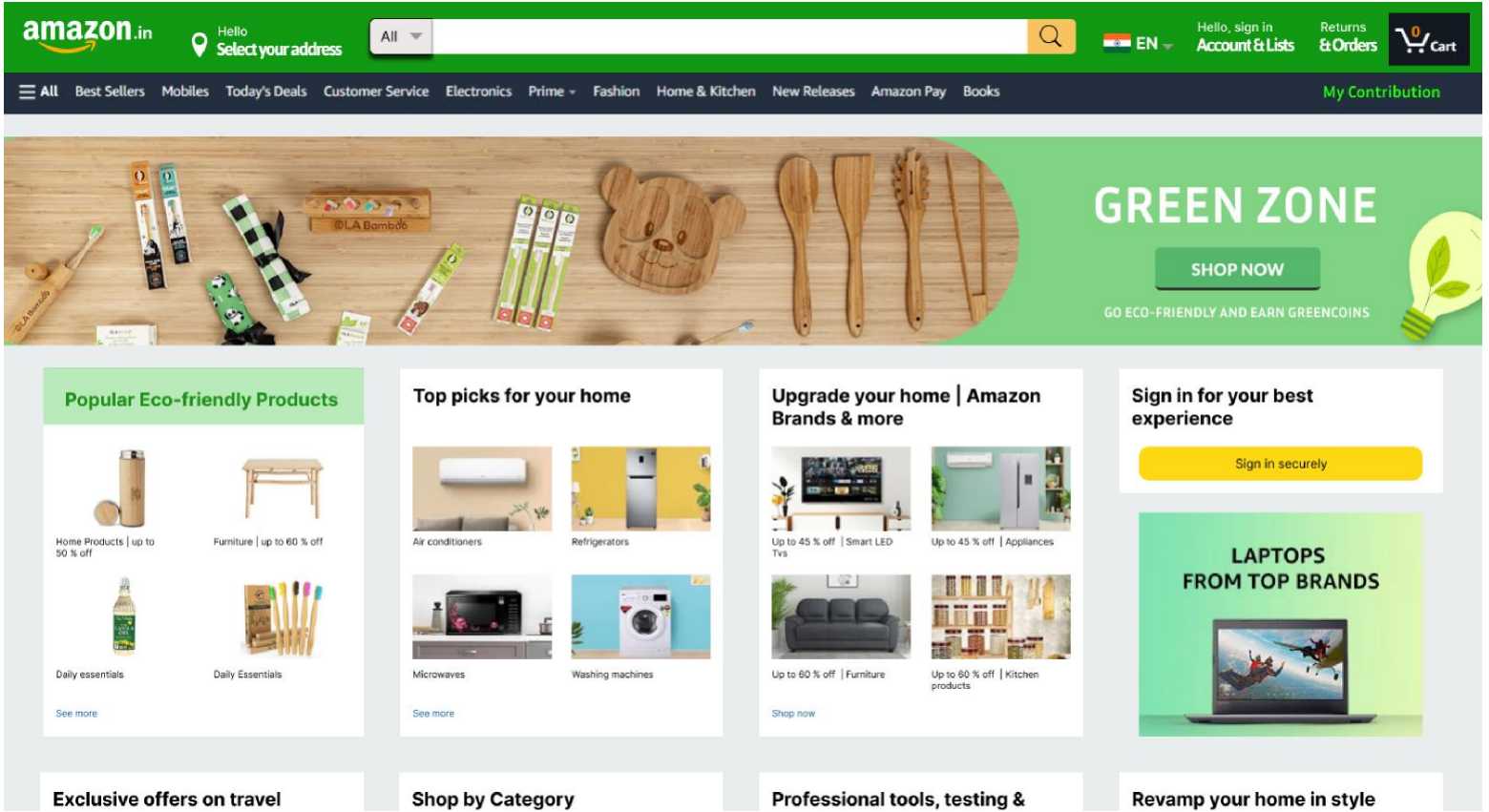


User Dashboard Design

- Users can see a visual representation of the impact their purchases have made.
- They can redeem their *Green Coins* to buy products or *Amazon services*.
- Users can also view the sectors where they have reduced pollution, expressed as a percentage. Monthly greenhouse gas reductions are also plotted.



Dedicated Green Portal Design



Sustainable Packaging Choices

- **User Checkout Options**
  - Standard Packaging (default)
  - Recycled Cardboard Box (+₹15)
  - Compostable Mailer
  - Pulp-Molded Insert (for fragile items)

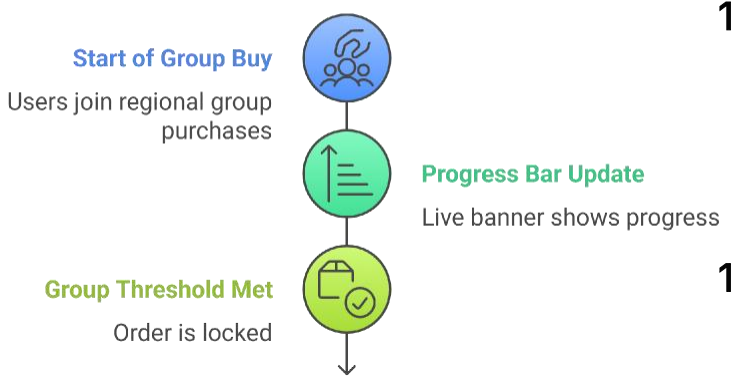
- **Displayed Metrics**
  - CO_2 emissions per packaging option
  - Extra cost (if any)
  - "You save X kg CO_2" badge etc.

- **ML Model – CO_2 Prediction**
  - Input Features: Cart weight, fragile items, delivery zone, packaging type
  - Model Used: Decision Tree Regressor (max depth = 4)
  - Workflow:
    1. Enumerate allowed packaging types
    2. *Predict total CO_2 = packaging CO_2 + shipping CO_2*
    3. Suggest lowest-emission option to user

- **Performance**
  - Inference time: ≤ 1 ms per option
  - Accuracy: *MAE < 0.1 kg CO_2* (validated)

Group Buying – Feature Overview & Fulfillment

1. **Feature Overview – "Group Buy" Mode**
   - Users can join regional group purchases of the same eco-product.
   - A live banner shows:
     - Progress bar (e.g., 3/5 joined, 24h left)
     - CO_2 + cost savings per user
   - Once the group threshold (e.g., 5 users) is met → order is locked.



1. **Fulfillment Flow**
   - A single bulk shipment is sent to a local hub or optimized delivery route.
   - Final deliveries made individually, reducing last-mile emissions.

1. **Benefits**
   - Fewer delivery trips → *Lower CO_2 emissions*
   - Reduced logistics cost per unit
   - Builds a sense of eco-community around shared actions

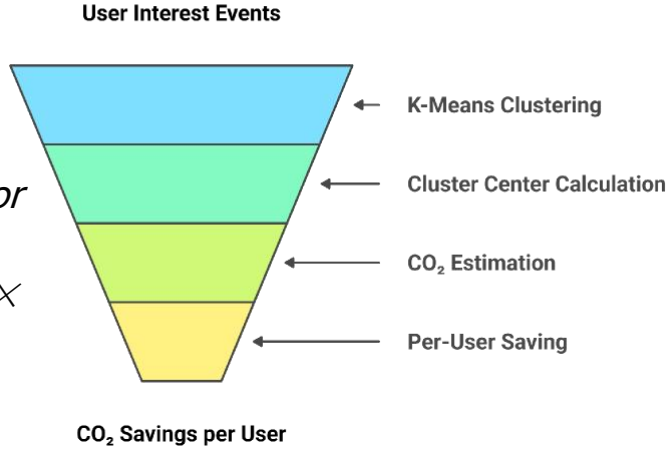
Group Buying – ML & Clustering Logic

1. **Clustering Logic**

We will run *K-Means* every 30 mins on 48-hour "interest" events

- *Inputs: (user_id, lat, lon, timestamp, ASIN)*
- *K = floor(total_users / target_group_size)*
- *Output: 3 top clusters of nearby users*

1. **CO_2 Saving Estimation**
   - *Individual CO_2 = dist(user → Fulfillment Center) × factor × avg_weight*
   - *Group CO_2 = dist(cluster_center → Fulfillment Center) × factor × avg_weight*
   - *Per-user Saving = Individual – Group CO_2*



1. **Smart UI Display**
   - On product page, show:
     - "Cluster 1: 8/10 joined – save 2.5 kg CO_2 each"
     - "Cluster 2: 5/10 joined – save 1.8 kg CO_2 each"
     - "Start new group – 48h left to fill 5 spots"

Business Relevance

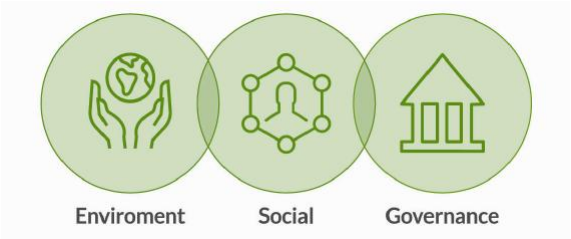
- **Reinforcing Amazon's Market Leadership**
  With Amazon already dominating the global e-commerce landscape, the integration of a "Greenovation Zone" strengthens its reputation as a forward-thinking leader in sustainability. It showcases Amazon not only as a commerce giant but as a brand with purpose – committed to driving responsible consumption.

- **Capturing the Premium Eco-Market**
  Consumer studies show that 55–66% of shoppers are willing to pay more for sustainable products. By offering verified green alternatives with transparency (eco-labels, CO_2 scores), Amazon can attract this premium, fast-growing segment and boost average order value and retention.

Business Relevance

Driving Loyalty Through Transparency & Impact

When users see their carbon savings, impact metrics, and eco-badges directly while shopping, they feel empowered. This transparency builds trust, strengthens emotional connection with the brand, and converts occasional buyers into loyal advocates for Amazon's green ecosystem.



Future-Proofing via ESG & Policy Alignment

Governments worldwide are pushing for greater transparency in sustainability claims. With built-in carbon tracking and verified eco-scores, Amazon can comply with upcoming ESG regulations while showcasing measurable climate action—helping meet net-zero goals and avoid greenwashing risks.

Future Scope & Modifications

Advanced Personalization: Incorporate user sustainability profiles (e.g. vegan, plastic-free), and use LLMs or semantic analysis to parse product descriptions for eco-signals. For instance, an AI "assistant" could answer user queries ("Which shampoo has lowest carbon footprint?").

Global Certification Standards: Establish partnerships with recognized global environmental certification organizations to ensure consistency and trustworthiness in eco-certifications for products listed in the Greenovation Zone.

Augmented Reality (AR): In a mobile app, allow users to scan product barcodes in-store to see instant eco-scores or find greener alternatives online.

Future Scope & Modifications

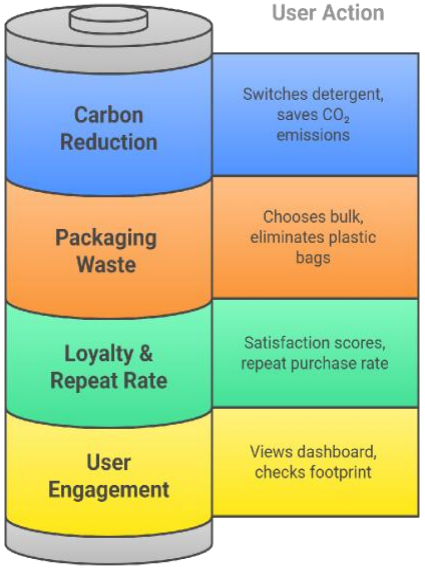
Supply-Chain Transparency with Blockchain:
We'll use a tamper-proof blockchain ledger to track every step of a product's journey—from raw material sourcing to final delivery. Each eco-certification and carbon footprint entry is immutably recorded, so shoppers can instantly verify a product's green credentials. This decentralized proof builds trust, prevents "greenwashing," and makes our sustainability claims 100% auditable.

ML-Powered Dynamic Pricing & Local Group Buys:
Leverage machine-learning algorithms to identify clusters of nearby shoppers with similar sustainable interests and purchasing patterns & collab these buyers form group. Then dynamically adjust discount tiers based on aggregated order volume. As the system unlocks deeper savings, motivating close-by green shoppers to pool their purchases and maximize both sustainability impact and cost benefits.

Success Metrics & Impact

We will measure both environmental and business impacts:

1. **Sales of Sustainable Products:** Track the increase in sales (or sales share) of eco-friendly items through our portal. (Studies show sustainable products' share in stores has risen steadily – e.g. ~22% in 2018 to a projected 25% by 2021)
2. **Carbon Reduction:** Calculate total CO_2 emissions saved per user and overall. For example, if a user switches to a biodegradable detergent, we quantify the emission difference. (The dashboard displays metrics like "500 kg CO_2 saved".)
3. **User Engagement:** Number of active users adopting green filters or viewing the dashboard. We can use metrics like % of shoppers who check their footprint or use the eco-recommendations.
4. **Packaging Waste:** Estimate the reduction in packaging waste (kg) by comparing standard vs. chosen options.
5. **Loyalty & Repeat Rate:** Monitor customer satisfaction scores, repeat purchase rate, and referral rates—especially among "green customer" segment.



Success Metrics & Impact

We can present these metrics in a table for clarity:

Metric	Impact KPI
Growth in Eco-Product Sales	+% increase in sustainable items sold (vs. baseline)
Carbon Footprint Reduction	Tons of CO_2e avoided per year
User Adoption	Tons of packaging material saved (per month)
Engagement with Green Features	Click-through-rate on eco-badges, dashboard views
Customer Satisfaction	Ratings/feedback from green shoppers

For example, industry data suggest that sustainability is driving growth: 50% of CPG sector growth (2013–2018) came from products marketed as sustainable. If we assume a similar uplift, our portal's curated focus could significantly boost sales in the green category.

Scalability & Marketplace Expansion

This solution is inherently scalable and extensible:

1. Broader Catalog: We can start with core categories (food, personal care) and later add others (apparel, electronics). Each new category just needs sustainability metadata added.

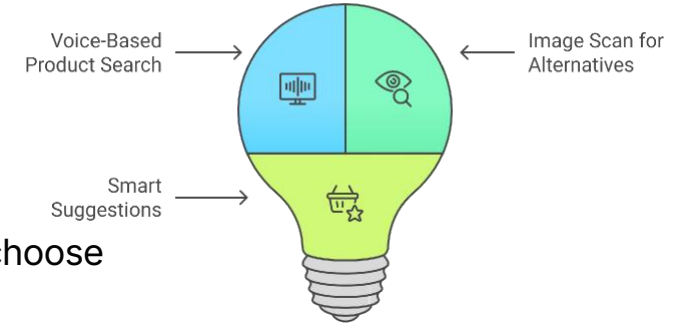
2. Multi-Vendor Integration: The portal could be extended to multiple merchants or a marketplace model. Third-party sellers could upload product green ratings to our platform.

3. Platform Scale: Using cloud services (AWS Elastic Beanstalk, EC2 auto-scaling, RDS or Aurora, etc.) ensures we can handle large user traffic. Microservice architecture allows independent scaling of recommendation engines and databases.

Scalability & Marketplace Expansion

4. Modular Architecture for Future Upgrades: Our system is built in parts (modules), so we can easily add new features without changing the whole setup. In the future, we can include:

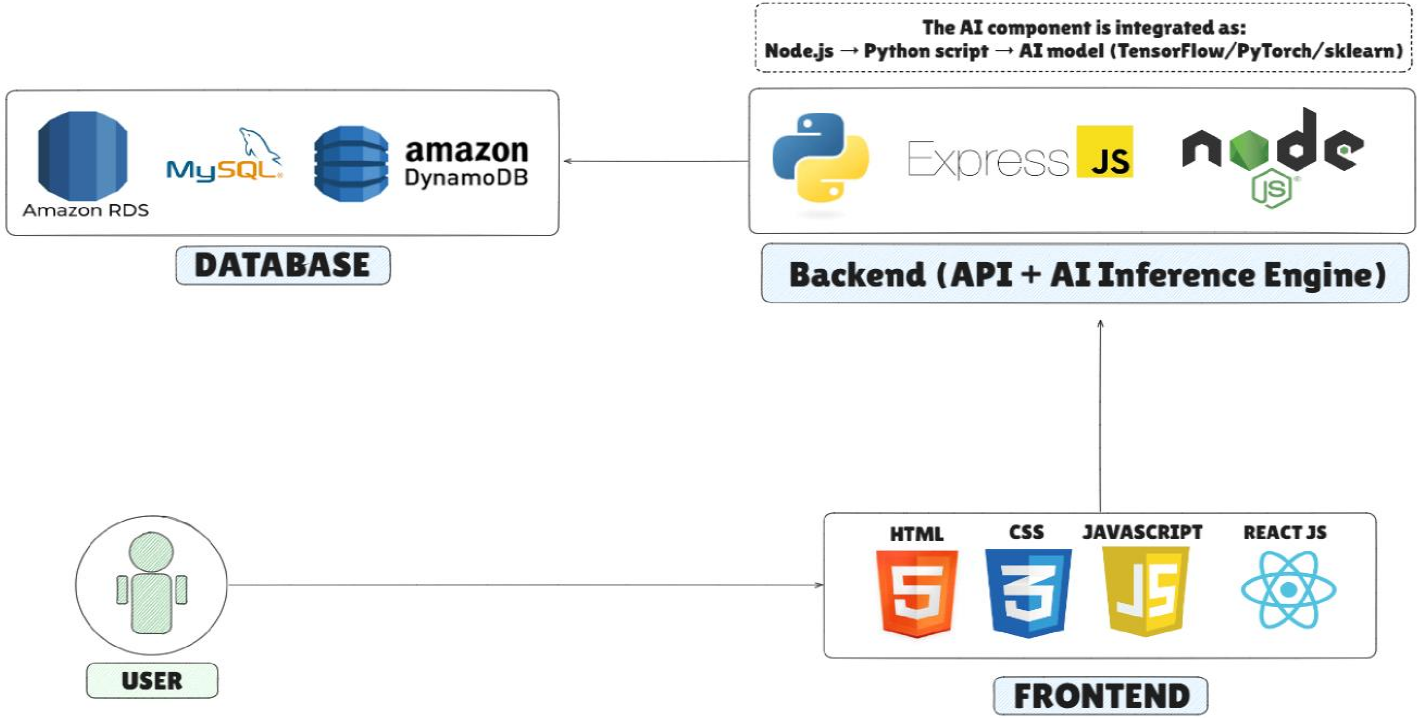
- Voice-based product search (e.g., “Show me eco-friendly shampoos”)
- Image scan to find eco-friendly alternatives
- Smart suggestions based on user habits (like reminders to choose green packaging)



5. Data Ecosystem: We can partner with eco-certification bodies (Fair Trade, USD Organic) to bulk-import product credentials. The system could also ingest real-tir updates (e.g. new carbon footprint studies) via APIs.

METHODOLOGIES

ARCHITECTURE DIAGRAM



Assumptions & Constraints

Aspect	Assumptions	Constraints/Challenges
Sustainability Data Access	Assume access to certified carbon and eco-data	Lack of standardization may need partnerships or audits.
User Adoption	Assume users care about eco-shopping	Risk of low engagement. Mitigation: gamify with badges, rewards, and impact dashboards.
AI Reliability	Assume accurate input data	Risk of biased or misleading recommendations if data quality is poor. So, rely on verified sources
Engagement with Green Features	Assume scalable backend for millions of products	High initial setup cost (data, ML, ETL). Worth it for long-term impact.

References

https://dokan.co/blog/497523/sustainable-shopping-trends/

https://www.pwc.com/gx/en/news-room/press-releases/2024/pwc-2024-voice-of-consumer-survey.html#:~:text=LONDON%2C%2015%20May%202024%20%E2%80%93,the%20Consumer%20Survey%2C%20published%20today

https://www.mckinsey.com/industries/consumer-packaged-goods/our-insights/consumers-care-about-sustainability-and-back-it-up-with-their-wallets

https://www.marinebiodiversity.ca/eco-friendly-consumers-10-eye-opening-statistics-how-you-can-join-the-green-revolution/#:~:text=Eco,the%20sustainable%20consumer%20movement%2C%20as

https://www.apu.apus.edu/area-of-study/business-and-management/resources/sustainable-ecommerce-strategies-for-eco-friendly-retail/

https://ssir.org/articles/entry/cultivating_the_green_consumer

amazon

Thank You!

Looking forward to hearing back from you