Algnition 2.0 Hackathon Challenge

Title: Hyper-Personalized Landing Page Generator Agent

🌟 Objective

Develop an AI-powered prototype that can dynamically generate hyper-personalized landing pages for *new or guest* users visiting an eCommerce website. The solution should intelligently recommend content and product modules based on inferred patterns of past user behavior, demographics, and transaction trends. Importantly, the prototype must also incorporate a strategy to address the **cold start problem** often encountered in recommendation engines.

✍️ Problem Statement

Today, digital commerce platforms are flooded with data, but struggle to personalize experiences for users with little to no history — such as first-time or anonymous visitors. Most recommendation systems rely heavily on behavioral or transactional history, and their effectiveness is limited in "cold start" scenarios where data about the user is not yet available.

Your task is to build a **Hyper-Personalized Landing Page Generator Agent** that can:

1. **Analyze patterns in historical data** to infer user interests and segment preferences.

2. **Recommend personalized landing page modules** (e.g., hero images, product carousels, CTAs) for incoming new users based on matching behavioral/demographic signals.

3. **Solve the cold start problem** by designing fallback strategies using available attributes like region, device type, traffic source, demographic cohort, and general trends.

In essence, the system should behave like an intelligent concierge for first-time visitors, tailoring the website experience based on what it has learned from past user behavior.

📊 Datasets Provided

1. User Activity Logs (GA4-style)

This dataset mimics how platforms like Google Analytics 4 collect user-level event data. Each row represents a specific event (e.g., session start, page view, item view, cart addition) performed by a visitor during their website interaction.

Key Columns Explained:

- user_pseudo_id: A unique identifier for a user across multiple events. Use this to group actions into sessions.
- event_name: Indicates the type of event. Examples include:
    - session_start: Marks the beginning of a user session.
    - page_view: The user viewed a specific webpage.
    - view_item, add_to_cart, purchase: Ecommerce-related interactions.
- category: The type of device used (e.g., mobile, desktop).
- city, region, country: Geographical location of the user.
- source, medium: Indicates how the user landed on the website (e.g., Facebook via PaidSocial).
- purchase_revenue, total_item_quantity: Populated only for transaction-related events.
- transaction_id: Transaction identifier, used to join with the purchase dataset.
- eventDate, eventTimestamp: Date and timestamp of the user event.
- gender, Age, income_group: Demographic information (may be inferred or user-declared).
- page_type: Section of the website visited (e.g., homepage, product page, collections, checkout).
- page_path: Full URL path of the visited page.

Sample User Flow Example: Let's say the same user_pseudo_id appears in multiple rows:

1. session_start on homepage
2. page_view of a collections page
3. view_item on a specific product
4. add_to_cart for that product
5. checkouts followed by a purchase

This sequence reflects a complete user journey from awareness to conversion and can be analyzed for personalization signals.

DataSet Link:-
https://drive.google.com/file/d/1G1EHGDsNctlKTusIuFKaYNIC0ycLuH4I/view?usp=drive_link

2. Transaction Dataset

This dataset contains detailed records of eCommerce transactions, with a focus on what users have actually purchased. It complements the activity logs by providing revenue and product-specific data.

Key Columns Explained:

- Transaction_ID: Unique identifier for each completed transaction. Use this to join with transaction_id in the activity log to link browsing behavior with final purchases.
- ItemID: A unique identifier for the purchased item. This can help in mapping user preferences or identifying frequently bought products.
- ItemName: The product name or title, which can be used for display or categorization.
- ItemBrand: Indicates the brand of the product. Useful for brand affinity analysis or brand-based recommendations.
- ItemCategory: High-level product classification (e.g., Accessories, Footwear, Apparel). This is critical for recommending relevant sections on the landing page.
- Item_purchase_quantity: The number of units of this specific item purchased in the transaction.
- Item_revenue: The total revenue earned from the sale of this item. This helps in identifying high-value products and revenue-driving categories.
- Date: The date on which the transaction occurred. Useful for seasonality or trend-based modeling.

How to Use This Dataset:

- Calculate **popular products** and **top-performing categories** by revenue or quantity.
- Identify **co-purchased products** or **category-level buying patterns**.
- Create a reference for **personalized product blocks** such as "Trending Now", "Top Picks by Category", or "Popular Among Similar Users".

Example Use Case: If a user viewed several items from the "Accessories" category but made no purchase, the agent can use this dataset to recommend bestsellers in the "Accessories" category based on historical purchase trends.

By combining behavioral indicators from the user activity log with actual transactional outcomes from this dataset, participants can design more intelligent, intent-aware landing pages.

DataSet Link:-
https://drive.google.com/file/d/1OxHOfTqL5nZW_IAyBB-JSImwyWMbVwk-/view?usp=drive_link

🔄 Expected Workflow

Participants are encouraged to follow a structured development pipeline:

1. Data Understanding & Preprocessing

- Join activity and transaction datasets using the transaction_id field.
- Construct meaningful user sessions and categorize them based on engagement type.

2. User Segmentation

- Group users by behavioral and demographic attributes:
    - Engagement levels: cart abandoners, frequent viewers, repeat purchasers
    - Demographics: age groups, gender, income brackets
    - Source: paid vs organic traffic, social channels
    - Geography: city, state, country

3. Cold Start Strategy Design

- Identify fallback logic using:
    - Similar users (via clustering or embedding techniques)
    - Default category trends by region, time of day, or device
    - Profile similarity-based logic using KNN, demographic filtering, or rules

4. Personalization Logic Development

- Define rules/models for what content to show on landing pages:
    - Hero banners based on inferred interest or top-trending categories
    - Product modules filtered by category popularity, age cohort interest
    - CTA modules adapted to stage of user (e.g., discover, explore, buy now)

5. Build the Prototype

- Optional but encouraged: Create a front-end UI mockup to showcase:
    - The real-time or inferred landing page layout
    - Visual sections aligned to segmentation logic

💳 Evaluation Criteria

Criteria	Description
Data Engineering & Strategy	How well data is preprocessed, joined, and leveraged to generate insights
Personalization Logic	Intelligence and relevance of the recommendations on the landing page
Cold Start Handling	Creativity and effectiveness of the fallback mechanism for new users
UI/UX (Optional)	Visualization clarity and how well the output prototype communicates the logic
Technical Execution	Code organization, clarity, use of AI/ML principles
Documentation	Clearly outlined assumptions, logic explanation, and model details

🚀 Submission Deliverables

Participants must submit the following:

- Source code: Python script, Jupyter Notebook, or Web App prototype
- A 5-slide deck that explains:
    1. Segmentation & data logic
    2. Personalization strategy
    3. Cold start solution
    4. Data architecture or ML pipeline (if applicable)
    5. Screenshots of output or UI demonstration
- (Optional) A live demo link or hosted prototype if applicable

🏆 Rewards

- ₹50,000 First Prize
- ₹25,000 Second Prize
- 3 Consolation Prizes for the Finalists
- Participation certificates for all entries

Entries

Can be Solo or Teams(Max 3 participants per team)

🔔 Deadline

All entries must be submitted by **30th June, 2025** through the mail to **AIgnition@netelixir.com**
Late entries will not be accepted.

❓ Need Support?

Reach out on the official AIgnition 2.0 WhatsApp Group or email us at **AIgnition@netelixir.com**

Let your imagination and data skills create the future of digital personalization. Good luck!