**4 Phase Interest model:**

Many students progress through school and early adulthood without developing meaningful interest in any academic or career domain. When a learner does not enter any phase of the interest‑development continuum Curiosity Activation, Engagement Sustainment, Personal Relevance Formation, or Passion‑Driven Mastery, they remain in a pre‑interest state where no subject feels engaging, relevant, or motivating. Over time, this lack of identified interest can lead to persistent academic disengagement, difficulty choosing educational pathways, low intrinsic motivation, and limited skill development. As students transition into higher education and adulthood, the absence of clear interests contributes to career uncertainty, reduced resilience, lower life satisfaction, and a tendency to drift rather than make intentional choices. This problem highlights the need for early detection of interest‑development gaps and structured support systems that help learners explore, discover, and cultivate meaningful interests across all stages of life.

**Why use a ML model?**

Using a machine‑learning model gives your interest‑detection system the ability to learn from real student behavior rather than relying only on fixed rules. As more responses are collected, the model begins to recognize subtle patterns such as mixed interests, emerging preferences, or phase‑specific engagement signals that a simple scoring system would miss. This makes predictions more accurate, more personalized, and more adaptive over time. ML also allows the test to evolve with different age groups, cultures, and learning styles, ultimately creating a smarter, more responsive system that grows better the more it is used.

**4 Phase of Interest (4PF):**

1. Curiosity Activation - Short bursts of attention sparked by novelty, surprise, or external cues.
2. Engagement Sustainment - Continued involvement supported by meaningful tasks, scaffolding, or relevance.
3. Personal Relevance Formation - Students begin to connect content to personal goals, values, and identity.
4. Passion-Driven Mastery - Deep, enduring interest with self-regulation, persistence, and proactive exploration.

**Age Groups:**

1. Primary (6–12 years):
   1. Curiosity Activation: Bright visuals, games, surprises, hands-on experiments.
   2. Engagement Sustainment: Structured tasks with immediate feedback, short challenges, group activities.
   3. Personal Relevance Formation: Linking topics to everyday life (pets, sports, family), encouraging storytelling.
   4. Passion-Driven Mastery: Early signs of “favorite subjects,” self-initiated projects (drawing, building, coding basics).
2. Secondary (13–18 years):
   1. Curiosity Activation: Real-world problems, debates, multimedia, social relevance.
   2. Engagement Sustainment: Longer projects, peer collaboration, gamified progress tracking.
   3. Personal Relevance Formation: Career exploration, elective choices, connecting learning to identity and values.
   4. Passion-Driven Mastery: Deep dives into chosen domains (science fairs, coding clubs, arts portfolios), self-directed study.
3. Higher Education (19–24 years):
   1. Curiosity Activation: Exposure to cutting-edge research, internships, interdisciplinary courses.
   2. Engagement Sustainment: Seminars, labs, collaborative research, mentorship.
   3. Personal Relevance Formation: Aligning studies with career goals, social impact, personal philosophy.
   4. Passion-Driven Mastery: Specialization, independent research, entrepreneurship, leadership in student organizations.
4. Lifelong Learners (25+):
   1. Curiosity Activation: New hobbies, workplace challenges, emerging technologies.
   2. Engagement Sustainment: Online courses, communities of practice, structured learning paths.
   3. Personal Relevance Formation: Skill-building for career advancement, personal enrichment, social contribution.
   4. Passion-Driven Mastery: Becoming mentors, thought leaders, or experts; integrating learning into lifestyle.

**Interest Domains:**

1. STEM (Science, Technology, Engineering, Mathematics)
2. Arts & Creative Expression
3. Business, Economics & Entrepreneurship
4. Health, Medicine & Life Sciences

**Dataset Structure:**

1. Metadata - Tag each question with age group, phase.
2. Answer Scoring - Map each option to a domain
3. Structured Responses - Record student answers with domain and phase for ML training.
4. Balanced Coverage - Equal representation across all domains and phases.
5. Age‑Appropriate Language - Simple for younger groups, abstract/career‑oriented for older learners.
6. Distractor Options - Neutral/unrelated choices to detect low interest or random answering.
7. Multi‑Domain Signals - Some options overlap domains to capture mixed interests.
8. Sample Size - Large, diverse responses per age group for generalization.
9. Synthetic + Real Data - Start synthetic, add real responses for accuracy.

Task: Generate dataset for each age group to find their interest

Dataset Resource: Real world uni and school data, Manual + LLM to generate and MCQs and refine them