TrustNet

An AI-driven Marketplace Trust & Safety Solution

HackOn with Amazon - Season 5

Submitted by: Nayan Sharma & Amil Mahajan
Team Name: SpongeCode & BugPants

Idea Submission Round | June 2025

TrustNet - An AI-driven solution for Counterfeit Detection and Review Authenticity

Problem Statement-

Due to the large scale of items sold on Amazon, it has become increasingly difficult to maintain customer trust because of counterfeit product listings and unreal reviews. It is common for customers to encounter fake branded products (e.g. "Air Jordans at too-good-to-be-true prices") and lots of positive reviews posted very quickly(e.g. 500+ five-star reviews appearing within an hour).

Scope of Innovation-

- Use Visual AI to detect counterfeit products by analyzing brand/logo anomalies in product images.

- Use LLM-based NLP for deep semantic analysis of product descriptions of products and review authenticity posted by users.

- Link sellers, listings, reviews and return events into a unified lifecycle scoring engine to detect fraud.

Who is the customer?

1 Anyone who shops for branded or premium products and wants genuine listings, reliable reviews and secure delivery.

Customer Pain Points

2
- ❑ Receiving counterfeit products because of highly rated listings.
- ❑ Falling for manipulated or AI-generated reviews.
- ❑ There is often no way to confirm a seller's identity before giving them your trust.

How TrustNet Helps

3 **TrustNet helps customers** by flagging suspect listings and reviews and previewing a 'trust score' before purchase and enables Amazon to act faster on fraudulent sellers.

Success Metrics

Impact of Trustnet

METRIC	CURRENT STATE (EST.)	AFTER TRUSTNET
Counterfeit Returns in Fashion	~14%	↓ 11.2% (−20%)
Fraudulent Listing Removal Time	~36 hrs	↓ 23 hrs (−35%)
Customer Trust Score	~70% (due to fakes, review spam)	↑ 80.5% (+15%) (more reliable listings, filtered reviews)

Reduced Returns

•When implemented, the return & refund rate for counterfeit products would go down by 20%.

Faster Takedowns

•There would be a 35% boost in how fast we took down fraudulent postings after they were detected.

Increased Trust

•Trust levels from customers would reportedly increase by 15% for items in high-risk sections (such as fashion and electronics).

Fewer Spam Reviews

•A decline in the ratio of spam reviews in the marketplace.

Impact

Loyalty and the Trust Customers Give

Raises trust in buyers thanks to genuine listings and reviews.

Lower refund and the cost of returning items

Brings certainty to our products by stopping counterfeit sales, decreasing returns and reducing items lost during shipping.

Protecting the reputation of your brand

Shields the reputable brands from being copied and replaced by fake products.

Keeping the Marketplace Free from Fraud

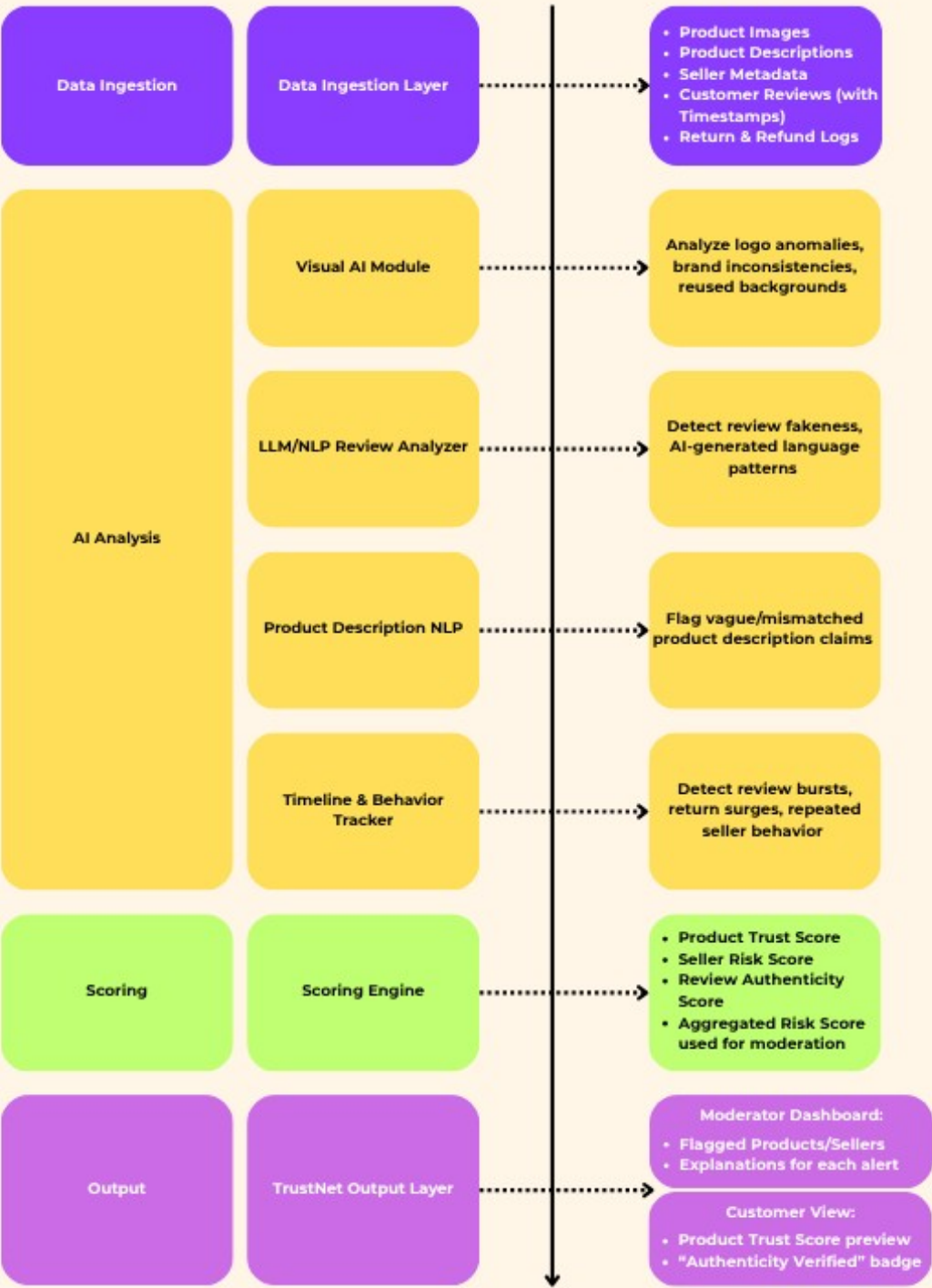AI is used to make sure the platform is reliable and protected in all its categories.

Architecture

Layers:-

1

Data Ingestion

Product images, bullet-point descriptions, seller profile, reviews and return logs.

2

AI Modules

Visual AI **(ResNet-50)**- To detect logo/product inconsistencies.

Text AI **(BERT or DistilBERT via SageMaker)**- To look for and flag anomalies in descriptions.

LLM **(Mistral-7B)**- For review classification + natural language explanation.

Timeline Engine **(Isolation Forest / LSTM)-** To detect review bursts and unnatural posting patterns.

3

Fraud Score Engine

Combines all types of risks and turns them into scores for each product and each seller.

4

TrustNet Dashboard

For customers- trust score preview.

For Amazon- check for high-risk sellers/products alerts.

TrustNet Processing Pipeline

- **Data Ingestion** → **Data Ingestion Layer** →
  - • Product Images
  - • Product Descriptions
  - • Seller Metadata
  - • Customer Reviews (with Timestamps)
  - • Return & Refund Logs

- **AI Analysis**
  - **Visual AI Module** → Analyze logo anomalies, brand inconsistencies, reused backgrounds
  - **LLM/NLP Review Analyzer** → Detect review fakeness, AI-generated language patterns
  - **Product Description NLP** → Flag vague/mismatched product description claims
  - **Timeline & Behavior Tracker** → Detect review bursts, return surges, repeated seller behavior

- **Scoring** → **Scoring Engine** →
  - • Product Trust Score
  - • Seller Risk Score
  - • Review Authenticity Score
  - • Aggregated Risk Score used for moderation

- **Output** → **TrustNet Output Layer** →
  - **Moderator Dashboard:**
    - • Flagged Products/Sellers
    - • Explanations for each alert
  - **Customer View:**
    - • Product Trust Score preview
    - • "Authenticity Verified" badge

AI-Stack Table

Task	Model / Tool	Why This Choice?
Image Anomaly Detection	ResNet-50	Fast, scalable
Review Auth. Classification	Mistral-7B Instruct	Open-source, LLM-based reasoning
Description Mismatch Check	DistilBERT via SageMaker	Compact, deployable on AWS
Burst Detection	Isolation Forest	Scalable, Interpretable anomaly detection

"TRUST SCORE COMPOSITION" BREAKDOWN
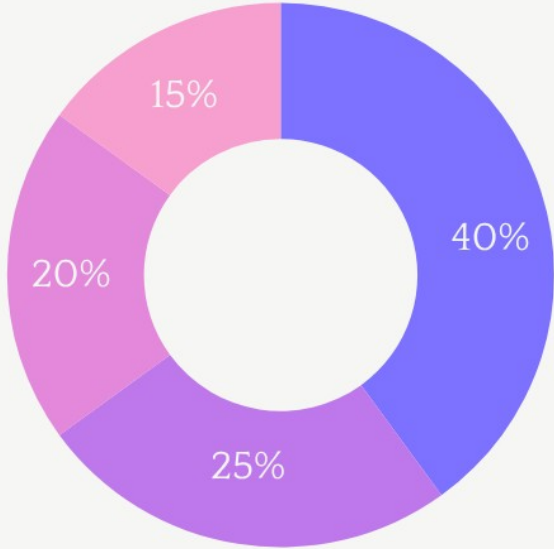
Final Product Trust Score =

- Visual authenticity (image AI) — 40%
- Review integrity (LLM) — 25%
- Behavior patterns (timeline) — 20%
- Text description mismatch — 15%

Scalability

🌐 Works across all Amazon marketplaces (India, US, Europe and so on).

🛡️ Adaptable for markets that have a high rate of counterfeit products(e.g. fashion, electronics, luxury).

⚙️ Modular engines- Review checker, Product verifier, Seller profiler.

Marketplace Integration

▢ Embedded at seller onboarding, listing approval, review monitoring and return assessment levels.

◎ APIs can be integrated within the company's internal tools for moderation and recognizing fraud.

Future Enhancements

Easy Scalability and adding APIs

We'll provide TrustNet APIs that anyone in Amazon can integrate into internal dashboards and fraud detection pipelines with little or no interruption or hassle.
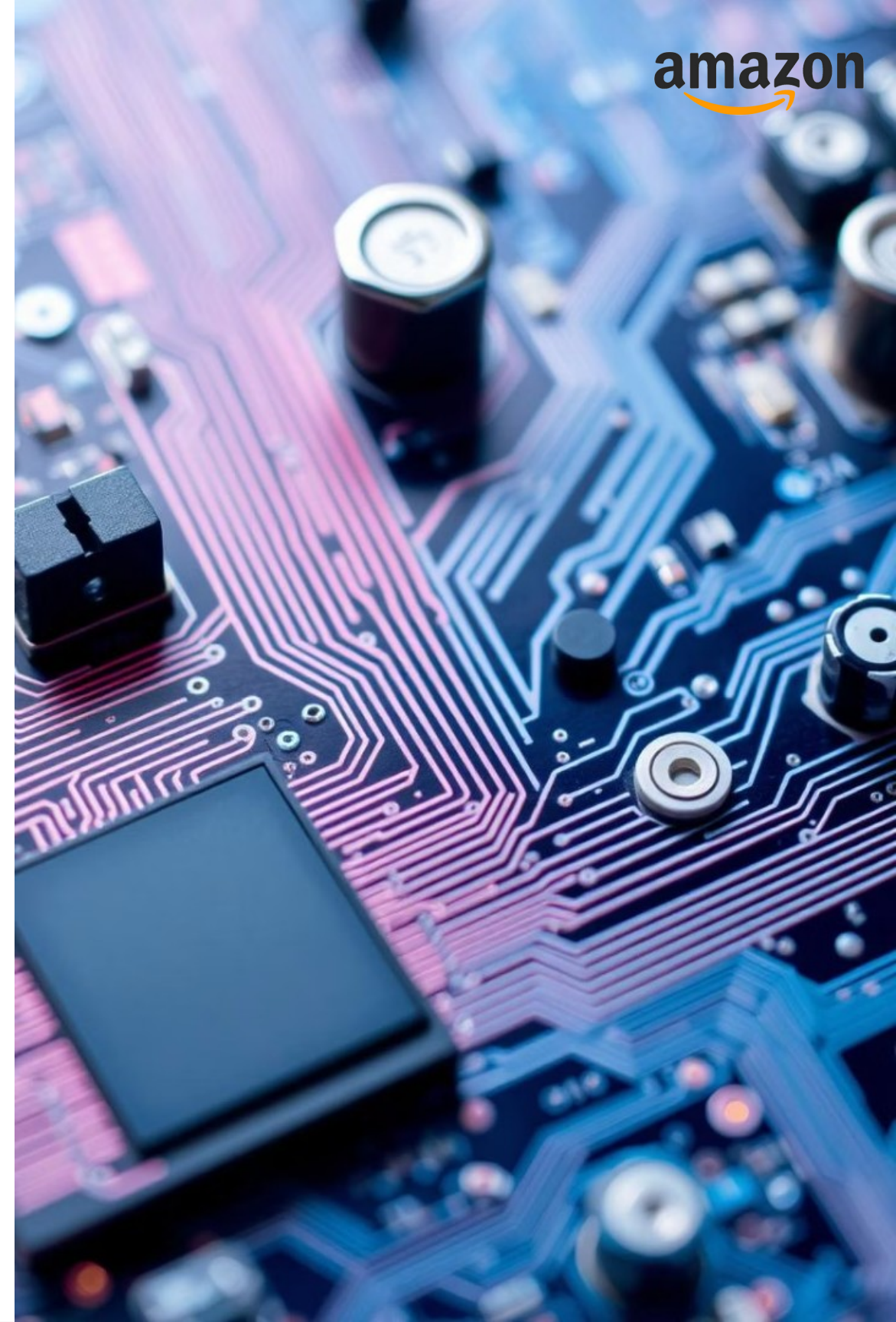
Customization of LLM

We plan to develop the context analysis of reviews and seller words by adjusting open-source LLMs such as Mistral based on Amazon's own information.

AI Modules Tailored for the Amazon Storage Environment

All three modules are designed to meet Amazon's rising need for reviewing new sellers, checking returned items and handling reviews of purchases.

Flagging Engine Mechanism

The system would provide a clear explanation along with every flag, as an example, "The logo does not match all images" or "Many reviews are nearly identical within 30 minutes."

Thank You