

BrainStation

Industry Project

Project:
Improve quality of Amazon reviews
based on Peer2Peer Validation Process

Team 2:

Juan Daza (SE)
Talha Pervaiz (SE)
Cristhian Lima (DS)
Mark Carlson (DS)

Project Managers

Jared Thomas (SE)
Rogelio Cuevas (DS)



Agenda

- ▶ Introduction
- ▶ Problem Statement
- ▶ Relevant Information
- ▶ Project Overview
- ▶ Solution Process
- ▶ Project Approach & Technology
- ▶ Amazon & Our Solution
- ▶ Summary & Future Vision
- ▶ Demo





Introduction

Did you know that Amazon hosts over a billion customer reviews? That's enough feedback to fill an entire library!

But how do we ensure that the information we're getting is reliable?

Imagine walking into a library where half of the books were written by someone who just glanced at the cover! That's how customers might feel when reading reviews that lack authenticity.

Our mission today is to transform the chaotic world of online reviews into something that actually helps customers make informed decisions.

Problem Statement

-
How might we...improve the quality of Amazon reviews while ensuring user's privacy preferences?

01

Authenticity and Verification

- Difficulty distinguishing genuine reviews from fake or incentivized feedback.

02

Inconsistent Review Quality

- Wide variation in review detail, usefulness, and relevance to the product.

03

Spam and Manipulation

- Presence of spam, excessive self-promotion, or review manipulation to influence product ratings.



Relevant information

-

Top 5 Amazon Consumer Review Statistics

01

Only 1-2% of Amazon buyers write reviews.

02

Fake reviews account for as much as 42% of Amazon's 750 million plus reviews.

03

82% of adults check reviews.

04

A one-star rating increase on Amazon can increase sales by 26%.

05

Product pages with customer reviews convert 3.5 times more frequently.

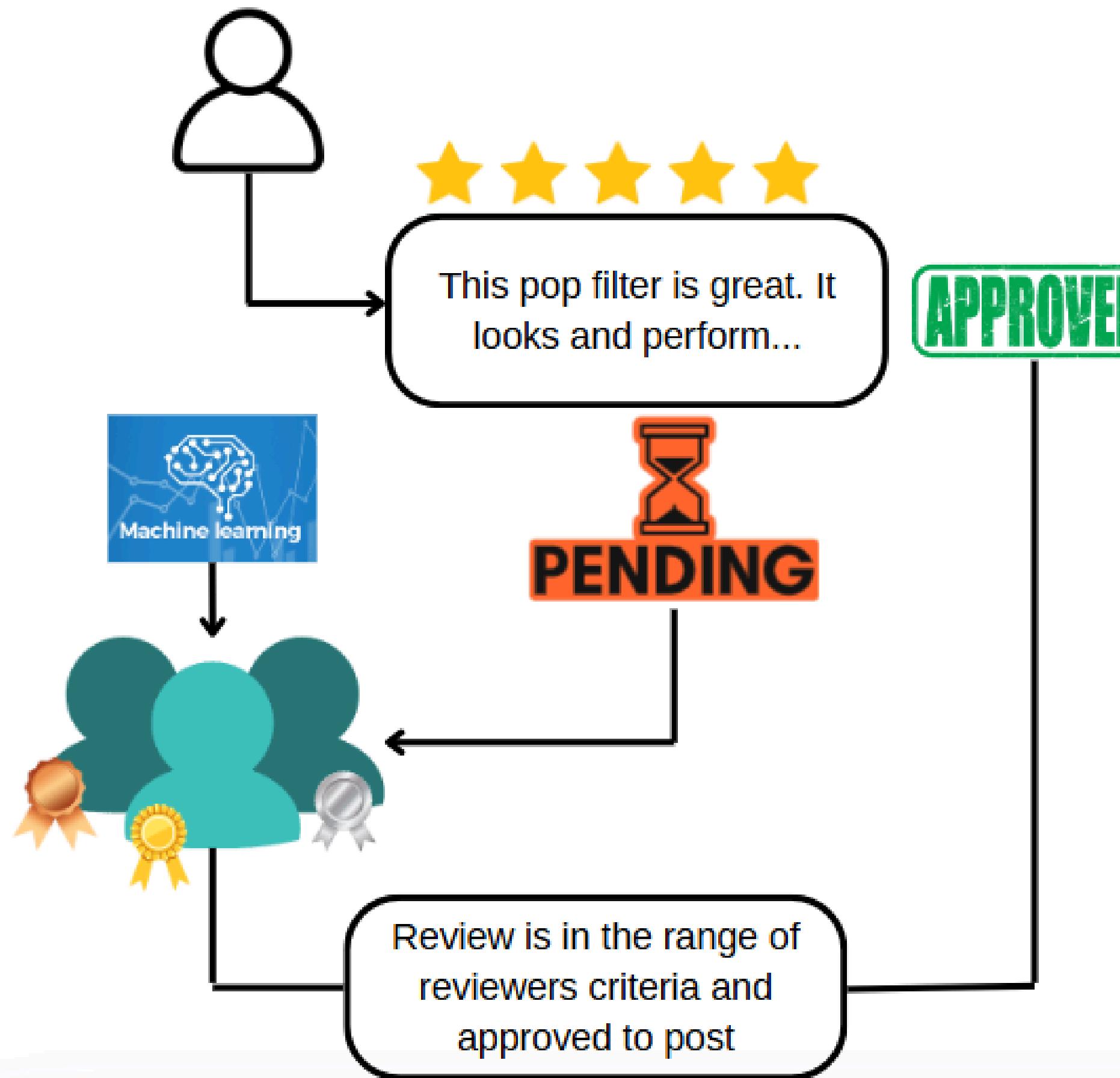
Project Overview

We are a group of data scientists and software engineers who have embarked on a journey to create a Peer2Peer validation system for Amazon reviews.

By assigning badges—bronze, silver, and gold—to reviewers based on their credibility, we aim to enhance the quality of feedback and restore consumer trust in the process.



Solution Process



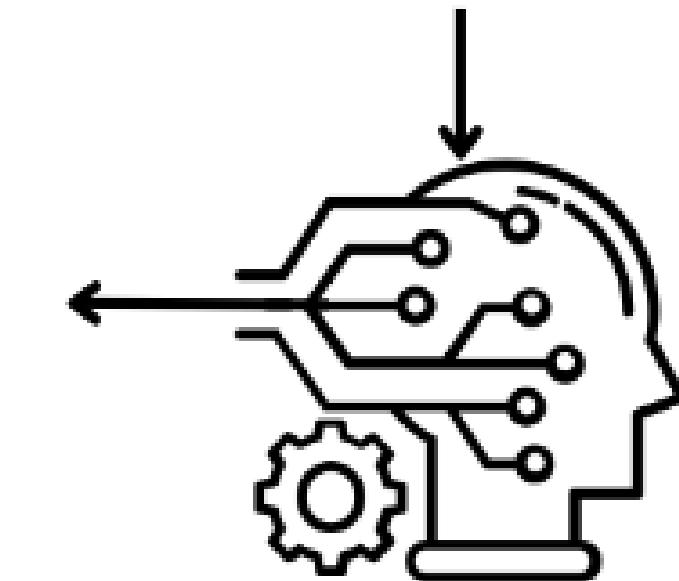
A new review is posted and keeps in pending status.

The review is sent to verified reviewers (based on badges)

Reviewers review and authenticate the new post based on their experiences with the product.

After approved by reviewers, the review is posted.
If not deeper analysis

ML model assign badges based on multiple criteria



Model for Identifying Reliable Reviewers

Objective: Identify and highlight the most reliable reviewers to create a trusted, peer-driven review validation system.

Solution Overview:

- **Classification Model:** A Decision Tree and NLP model identifies reviewers' reliability levels, categorizing them as Bronze, Silver, or Gold. This allows us to designate high-quality reviewers for validation roles.
- **Web Application Integration:** The model is seamlessly integrated into a web app:
 - **Badge Display:** Reviewer badges (Bronze, Silver, Gold) are visible in the app, showcasing reviewer reliability at a glance.
 - **Real-Time Validation:** Each new review is assessed by the model, which assigns validation responsibility to the highest-rated reviewers.

Business Impact:

- Trust & Transparency: Elevates the credibility of reviews by ensuring that only top-rated reviewers validate new entries.
- Automated Reliability Assessment: Reduces manual oversight and enhances user trust through an unbiased, data-driven approach.

Project Approach & Technology

Data Science Approach:

- Amazon review dataset
- Model that assign badges (Bronze, Silver, Gold) and identify potential reviewers. Main features to evaluate:
 - helpfulness_ratio: Proportion of helpful votes received.
 - total_votes: Total votes on each review.
 - average_helpfulness_score: Average helpfulness score for each reviewer.
 - overall: Review rating (1 to 5).
 - is_reliable: Target variable indicating reliable (1) and non-reliable (0) reviewers.
 - reviewText: To enhance classification, future iterations will apply text analysis on review content.



Software Engineering Approach:

- Developed a prototype displaying badges on profiles and reviews using React (front end) and Node.js (back end).
- Planned API integration to update badges, consuming model in real time.



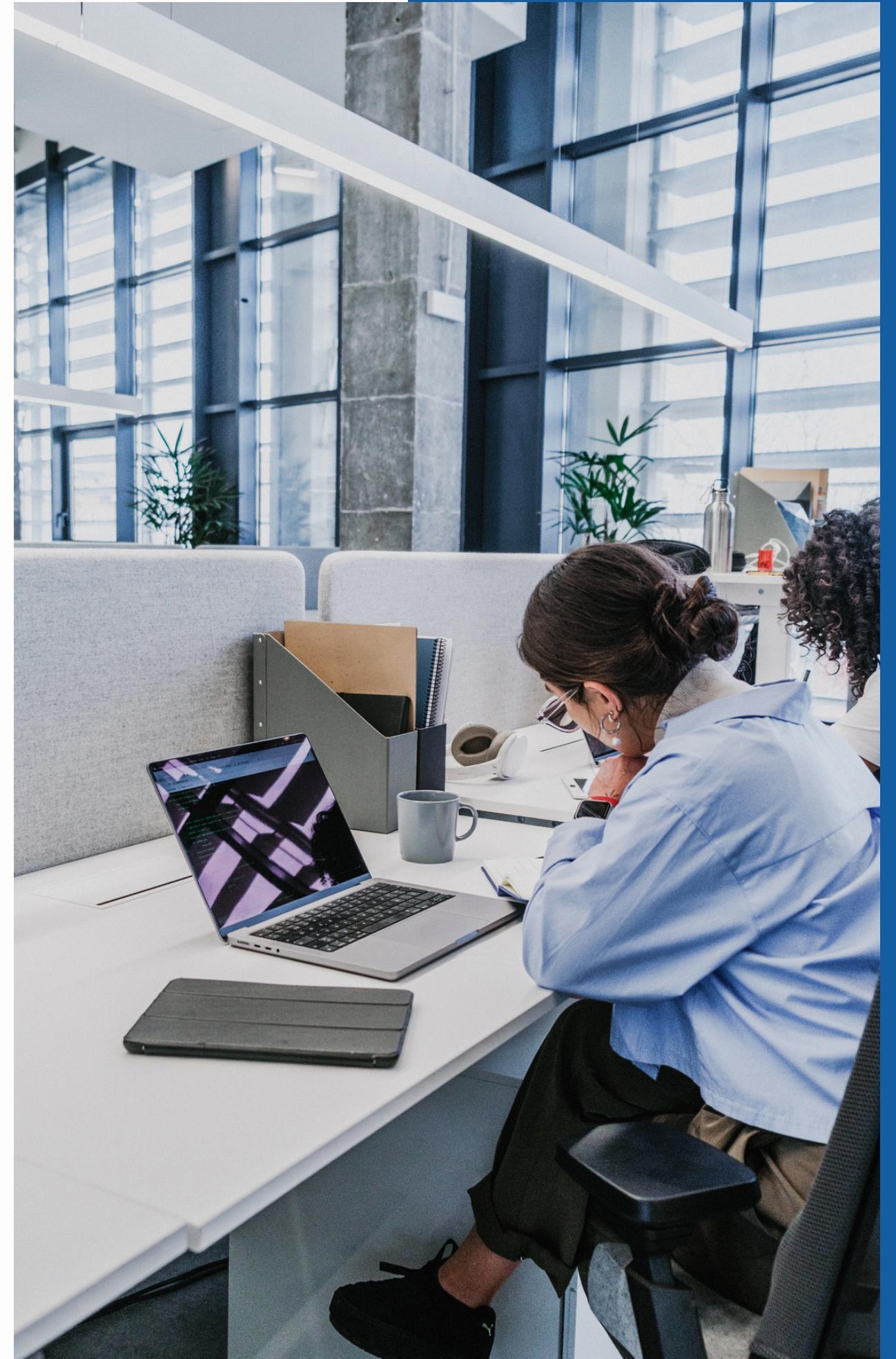
Amazon & Our Solution

Amazon Vine Program

- Purpose: Helps customers make informed decisions by inviting trusted reviewers to try products for free.

How Our Solution Enhances Vine

- Badge System: Adds credibility and provides accurate assessments of reviews, complementing Vine program the way participants are selected.



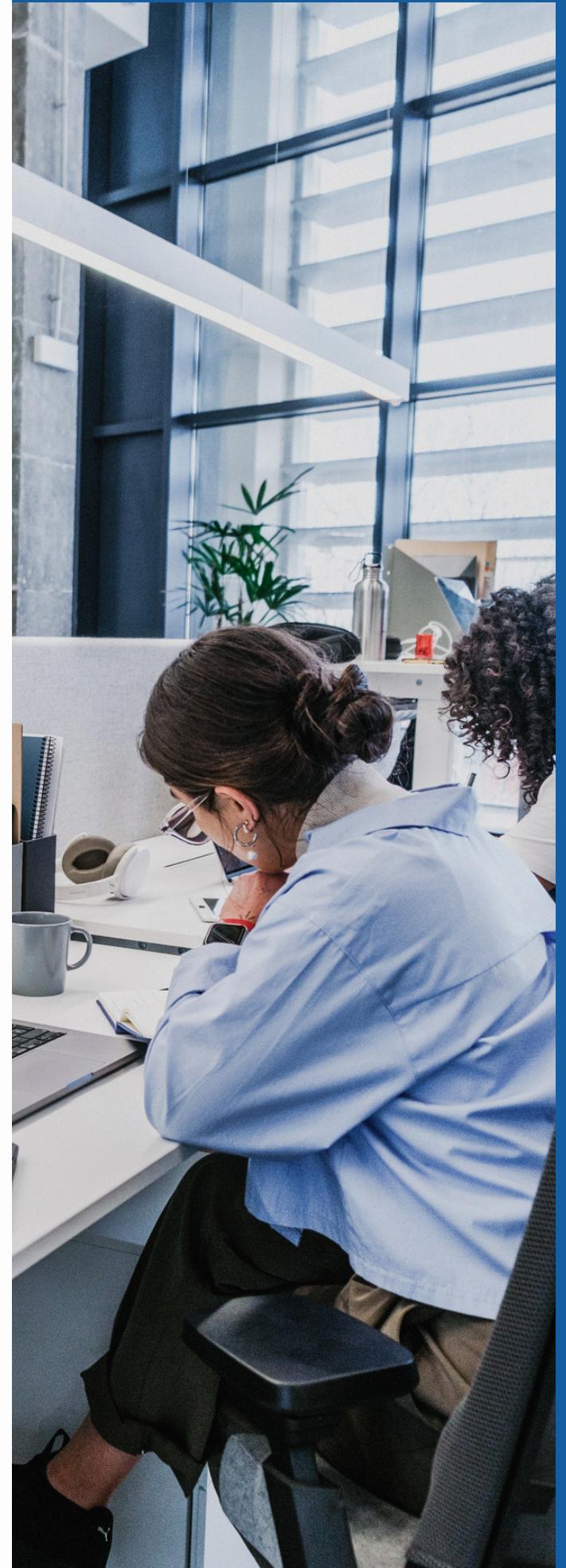
Summary & Future Vision

Recap of Our Approach

- Introduced a Peer2Peer validation system to improve Amazon reviews' quality and reliability.
- Enhanced user experience by awarding Bronze, Silver, and Gold badges based on review helpfulness, reliability, and recent activity.

What Makes Us Unique

- Data-Driven Insight: Leveraged sentiment analysis and behavioral data to identify trustworthy reviewers.
- Cross-Disciplinary Team: Combined skills from Data Science and Software Engineering to create a full-feature solution with a strong technical foundation and seamless user interface.

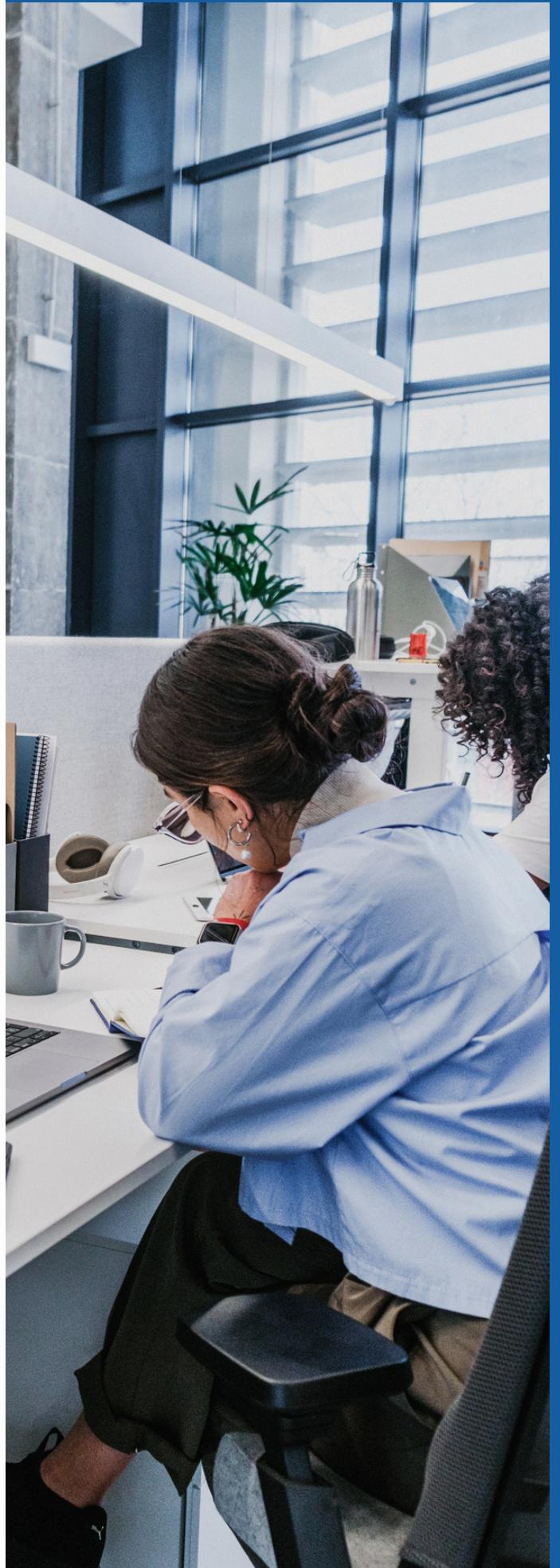


Summary & Future Vision

Future Goals & Expansion

- Stretch Goal: Incorporate sentiment analysis to further refine badge assignment, aligning reviewer categories more closely with user sentiment.
- Integration with Amazon Vine: Our system can add value to Amazon's Vine program by enhancing review authenticity and filtering, strengthening trust in product reviews.

DEMO →



THANK YOU!

