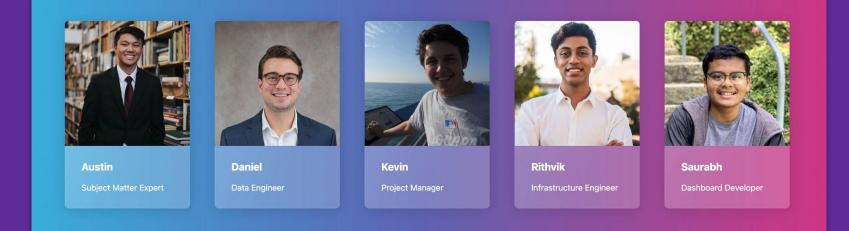
E-Commerce Price Prophet

Capstone Final Presentation

Austin Chen, Daniel Myers, Kevin Cahillane, Rithvik Sunku, Saurabh Narain



Our Team



Shouldn't e-commerce be easier?

WHY?

Pricing products is time-consuming

Platforms prioritize profits over fairness

WHAT?

Buyers face confusion due to price variations

An unbiased tool could improve fairness



\$3 trillion

Estimated revenue from all global e-commerce transactions in 2023



E-Commerce

From a Domain Expert

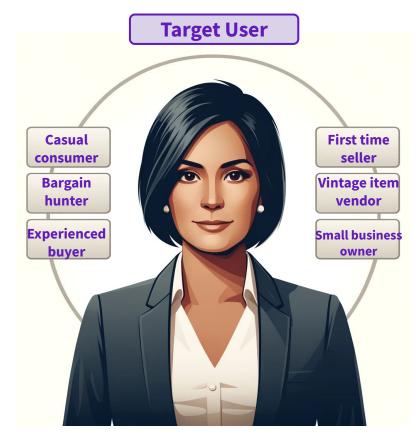


5:51 6	•	? 100
×	Sell an item	List
Add up to 12 photos		
Description		
Tissot rock wat	tch	18/80
Describe your	item (5+ words)	
		0/1000
Add up to 5 ho	shtags (Optional)	
		0/5
Details		
Category	Men > Men's accessories > Watches	>
Brand	Tissot	2
Condition	Good	>
Color	Select color (optional)	>
Delivery		

What does the ideal experience look like for users?

Price Prediction	Range		
Based on the information you've provided, we believe your product should be priced at around \$52.65 with a lower range from \$22.11 to an upper range of \$75.82			
\$ 22	\$ 53	\$ 76	
•	=	THE OWNER OF THE OWNER OWNER OF THE OWNER OWNE	

Buyers and sellers benefit from transparent pricing



Trust: Transparent pricing builds trust, fostering confidence in transactions and e-commerce platforms

Market Efficiency: Clear and consistent pricing reduces confusion, streamlining the market for everyone

Fair Competition: Sellers at all experience levels can compete more fairly, offering better prices and quality, benefiting consumers

Our data

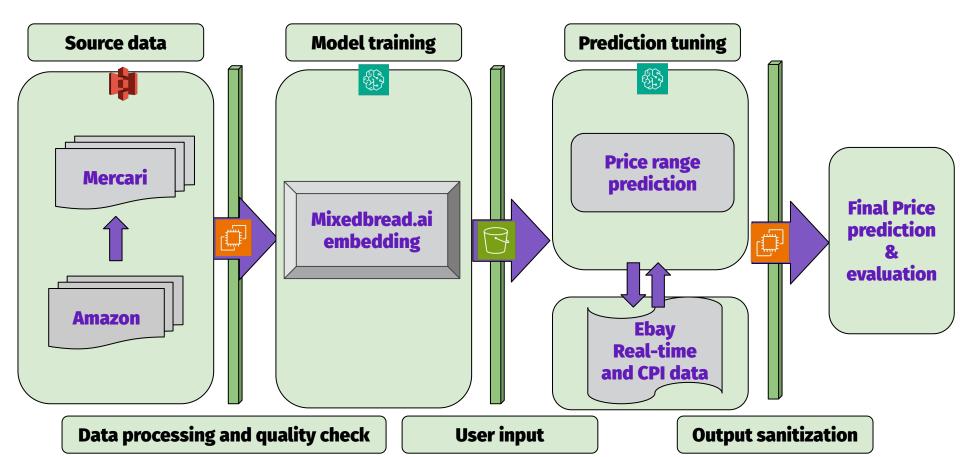




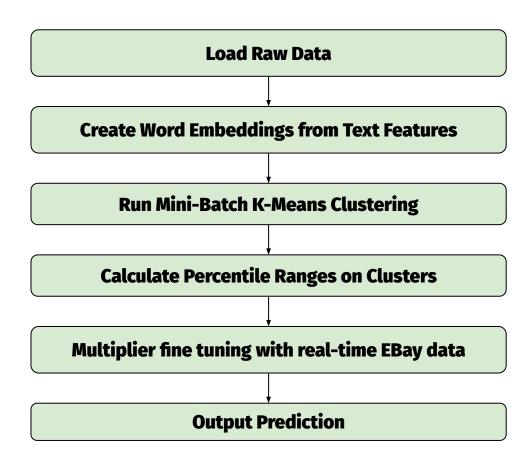


Dataset	Mercari Price Prediction (Primary)	Amazon Review (Supplementary)	eBay current listings (Secondary)
Observations	More than 700k	240 million in original dataset (100k most recent listings utilized)	120k+ observations recorded based on trained cluster characteristics
Fields	Item title, brand, category, condition, final sales price	Item title, brand, category, condition, final sales price	Item title, brand, category, condition, current listing price
Source	Kaggle	UCSD data repository	eBay Finding API
Limitations:	Historical listings collected from 2018-2019	Most recent listings recorded in 2018	Current listings; no final sales price

Data Architecture Diagram

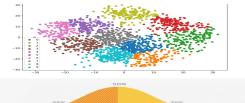


Model Pipeline Overview







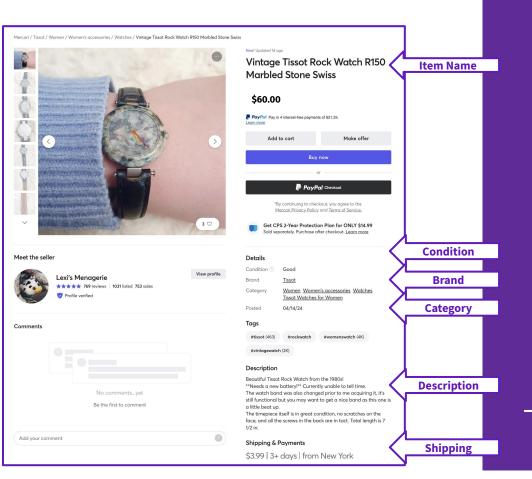






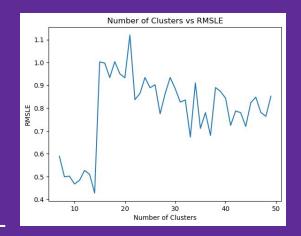


K-Means Clustering



Minibatch KMeans algorithm on training data, use cluster assignment to have 30% -70% quantile ranges for each cluster

Convert item name and description into vectors using MixedBread AI



Multiplier Model for Fine-tuning Price Ranges



This method allows us to create a unique pricing analysis per item

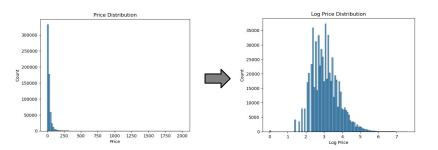
MVP Live Demo

A Mercari-based pricing dashboard that compares similar postings and recommends a price range for sellers

Results and Evaluation

We used <u>root mean square</u> **logarithmic** <u>error</u> (RMSLE) as our error metric in evaluating model performance and improvement

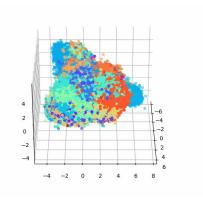
In the end, we achieved an RMSLE of 0.42 (cross-validated)



Transforming Our Data to Account for Heavy Right Skew

$$RMSLE = \sqrt{rac{1}{N}\sum_{i=1}^{N} \left(\log(y_i+1) - \log(\hat{y}_i+1)
ight)^2}$$

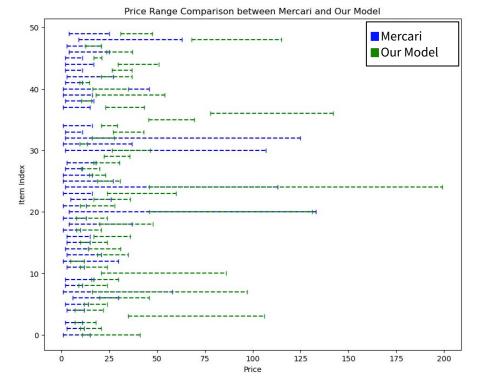
New Error Function to Account for Skewed Distribution

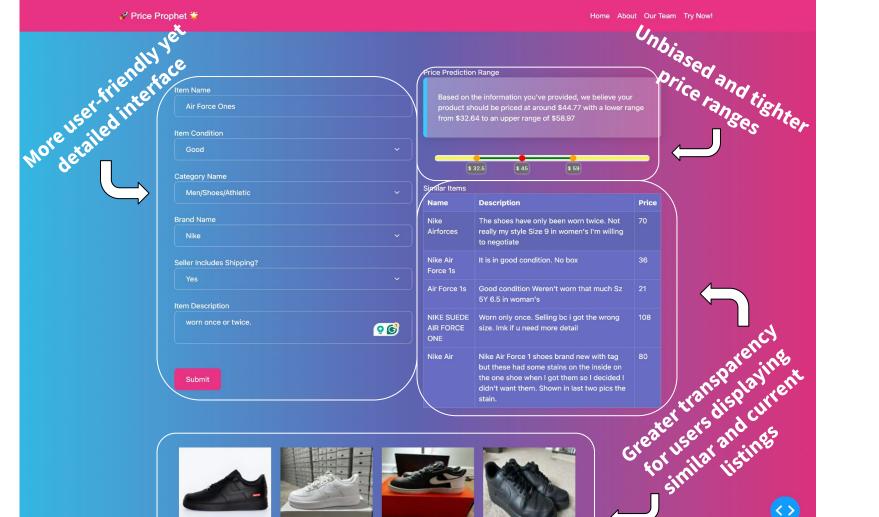


Ebay Data Vectorized by Our Model Projected in 3D Space

Benchmarking Against Mercari's Suggested Prices

To further validate and compare our model, we took an unseen sample from our data and used Mercari's suggested pricing tool and our tool.



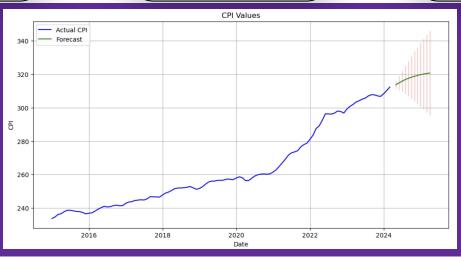




<>

Future Developments





Mission Statement

To optimize e-commerce efficiency and promote price transparency for all.

