



Alpha Presentation

Amazon Review Confidence Tool

The Capstone Experience

Amazon

Collin Cole

Nikita Gupta

Cameron Hurley

Ashu Kher

Dylan Mccarroll

Ethan Strain

Department of Computer Science and Engineering
Michigan State University



*From Students...
...to Professionals*

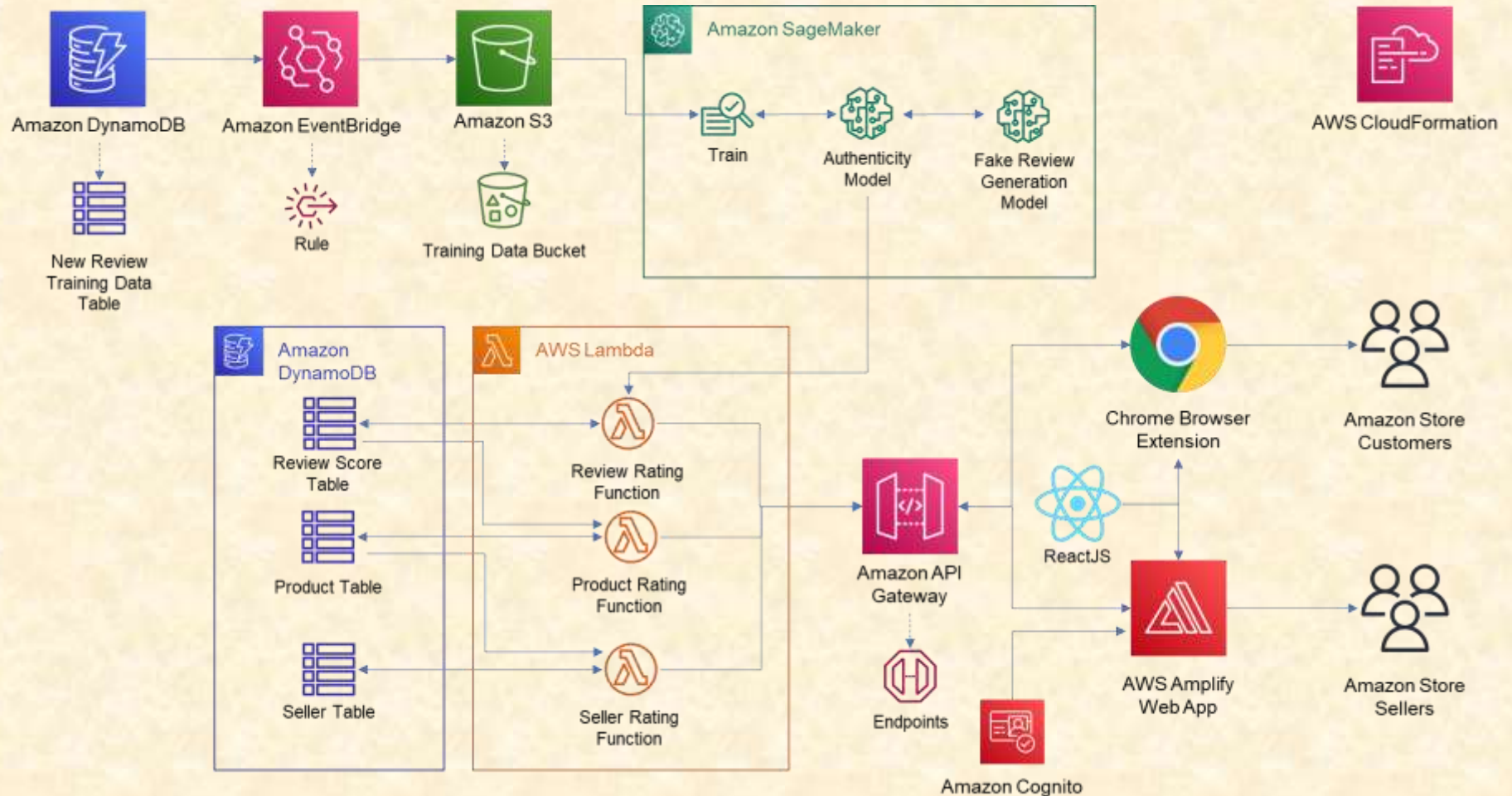
Fall 2022

Project Overview

- Review Authenticity Tool
 - Predict if an Amazon product review is genuine for sellers & customers alike
 - Visualize confidence score with easy to understand graphics
- Easily Accessible Data
 - Browser extension to display data on-site on Amazon products
 - Web page to provide further information and additional features



System Architecture



Home Page

The screenshot displays the Amazon homepage layout. At the top is a dark navigation bar with the Amazon logo, a search bar, and a home icon. Below this is a promotional banner for a browser extension. The main content area is divided into two sections: 'Saved Items' and 'Featured Items'. Each item in these sections includes a product image, title, price, and a star rating.

*** Got our Product Review browser extension? Install it [HERE](#) ***

Saved Items

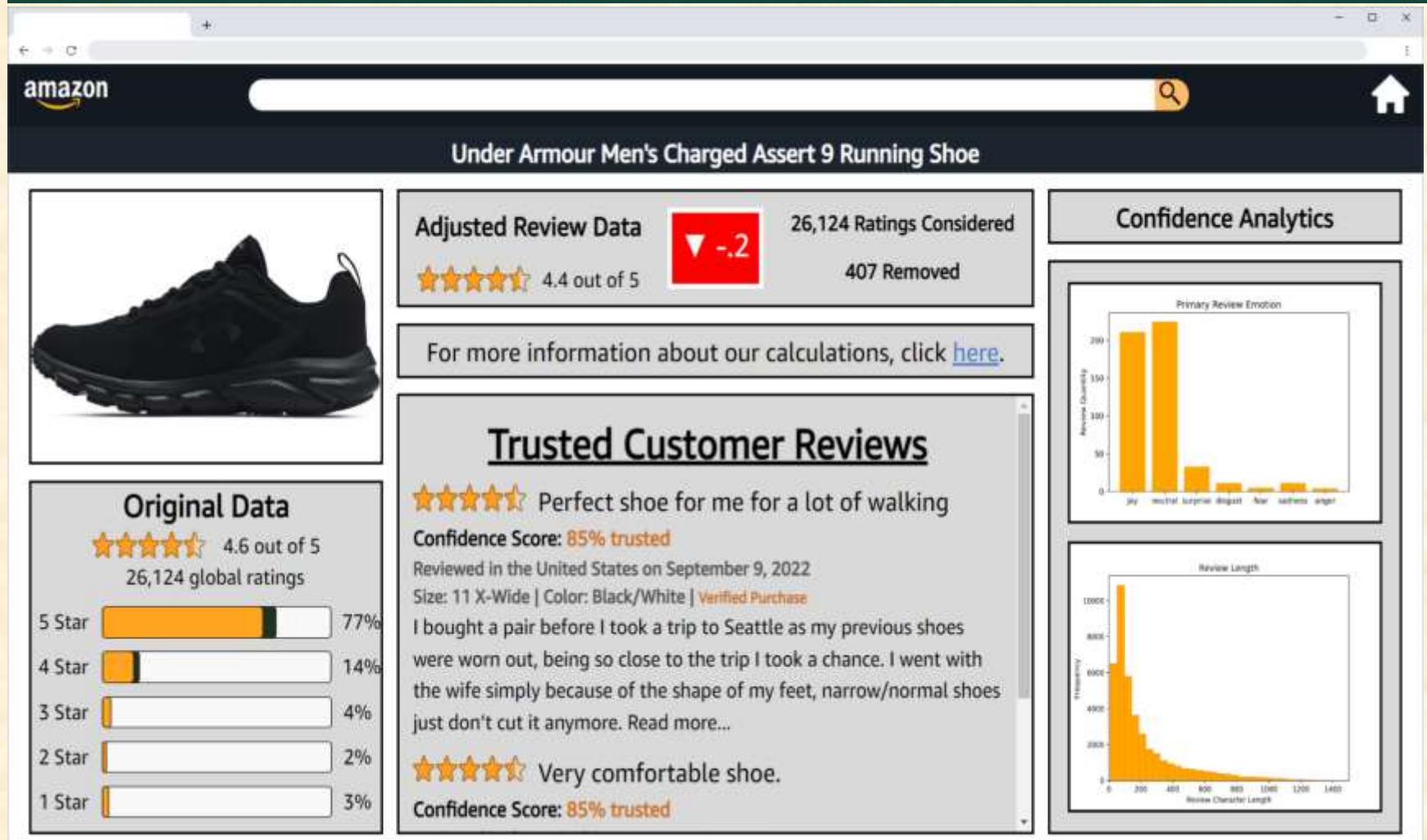
Product Image	Product Name	Price	Rating
	PlayStation 5 Console	\$499.99	★★★★★
	Nintendo Switch with Neon Blue and Neon Red Joy-Con	\$299.99	★★★★★
	Xbox Series X	\$499.99	★★★★★

Featured Items

Product Image	Product Name	Price	Rating
	Apple AirPods Pro (2nd Generation) Wireless Earbuds with MagSafe Charging Case. Active Noise Cancelling, Personalized Spatial Audio, Customizable Fit, Bluetooth	\$239.99	★★★★★
	(Pack of 2) Replaced Remote Control for Roku TV, Compatible for TCL Roku/Hisense Roku/Onn Roku/Sharp Roku/Element Roku/Westinghouse	\$11.57	★★★★★
	10Ft Long Surge Protector Power Strip, Extension Cord with 8 AC Outlets and 4 USB Ports(1 USB C), 3-Side Outlet Extender Strip, Flat Plug, Wall Mount, Dorm Home	\$19.86	★★★★★



Product Page



Browser Extension (On Amazon)

The screenshot shows the Amazon product page for the Under Armour Men's Charged Assert 9 Running Shoe. The product is a black sneaker with a thick sole. The page includes a navigation bar with the Amazon logo, delivery location (Grand Rapids 49514), and search results. The product title is 'Under Armour Men's Charged Assert 9 Running Shoe'. The 'Adjusted Reviews' overlay is positioned on the right side of the product image. It displays a star rating breakdown and an adjusted rating of 4.5 out of 5. Below the rating, it states '26,124 Original Ratings' and '407 Removed'. A link is provided for more information and metrics.

Star Rating	Percentage
5 Star	77%
4 Star	14%
3 Star	4%
2 Star	2%
1 Star	3%

Adjusted Rating
★★★★★
4.5 out of 5
26,124 Original Ratings
407 Removed

For more information and metrics, visit our web application [here](#).

Color: Black (002)/Black

Size Chart

Product Details



Fake Review Generation

```
response = client.invoke_endpoint(EndpointName="ElectronicsGenModel",
                                   ContentType='application/json',
                                   Body=data_bytes)

# load the response into clean variable
result = json.loads(response['Body'].read().decode())

# clear out the newline chars from output
generated_reviews = result[0]["generated_text"].split("\n")

for review in generated_reviews:
    print(review)
```

Great camera quality! I found it quite nice and accurate. If the other review had been the same then I could probably have said to myself: there is not a better camera in this price range.

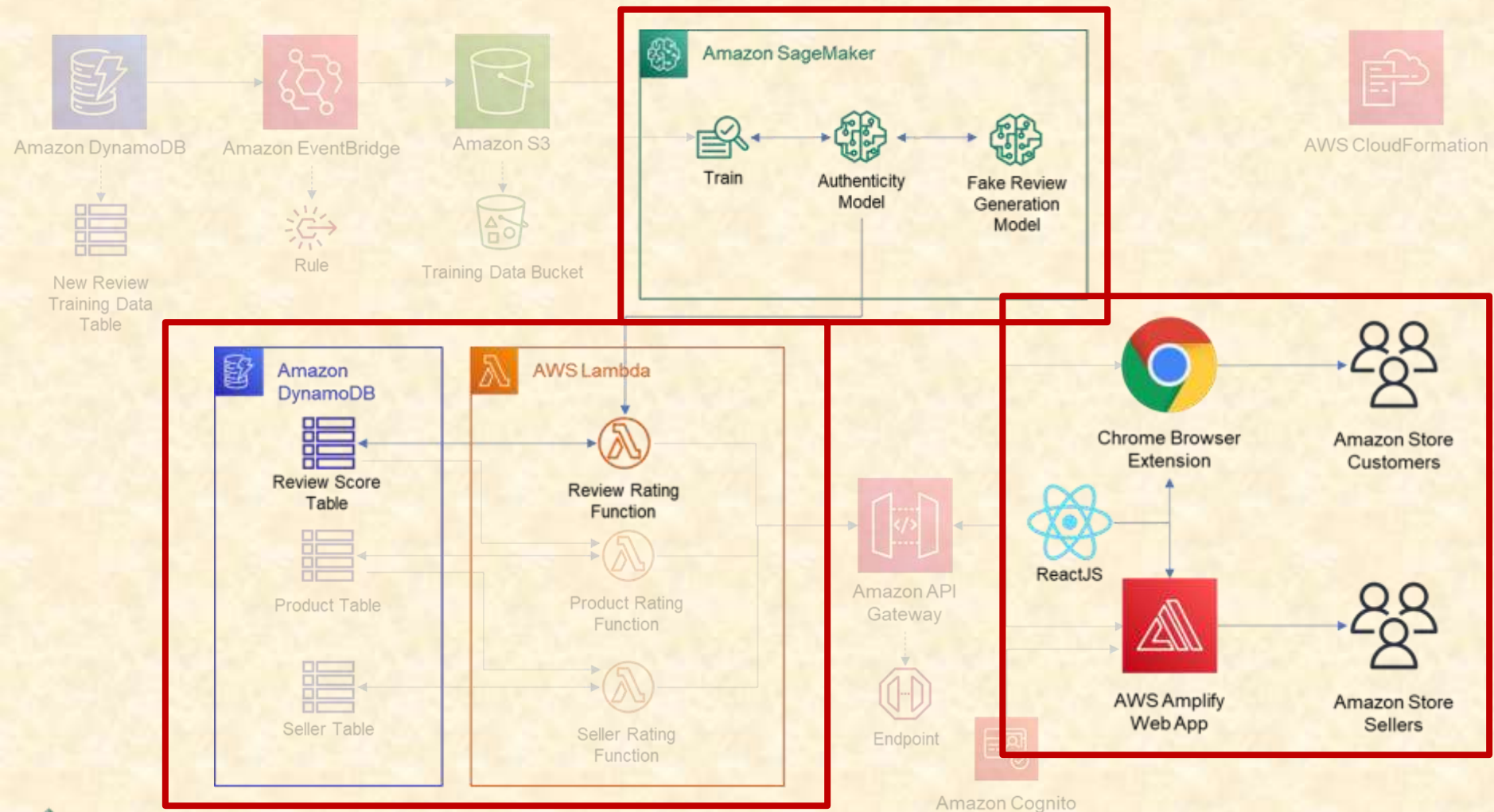
Rated 5 out of 5 by daniel k from Very simple and fast Very simple and fast

Rated 5 out of 5 by DZB from Great price For what it is I bought this camera to look at some vintage film. It has a great quality and the sound is clear in quality. I want to give it a shot and to say thank you all... for this camera it was a great deal. It was very good price.

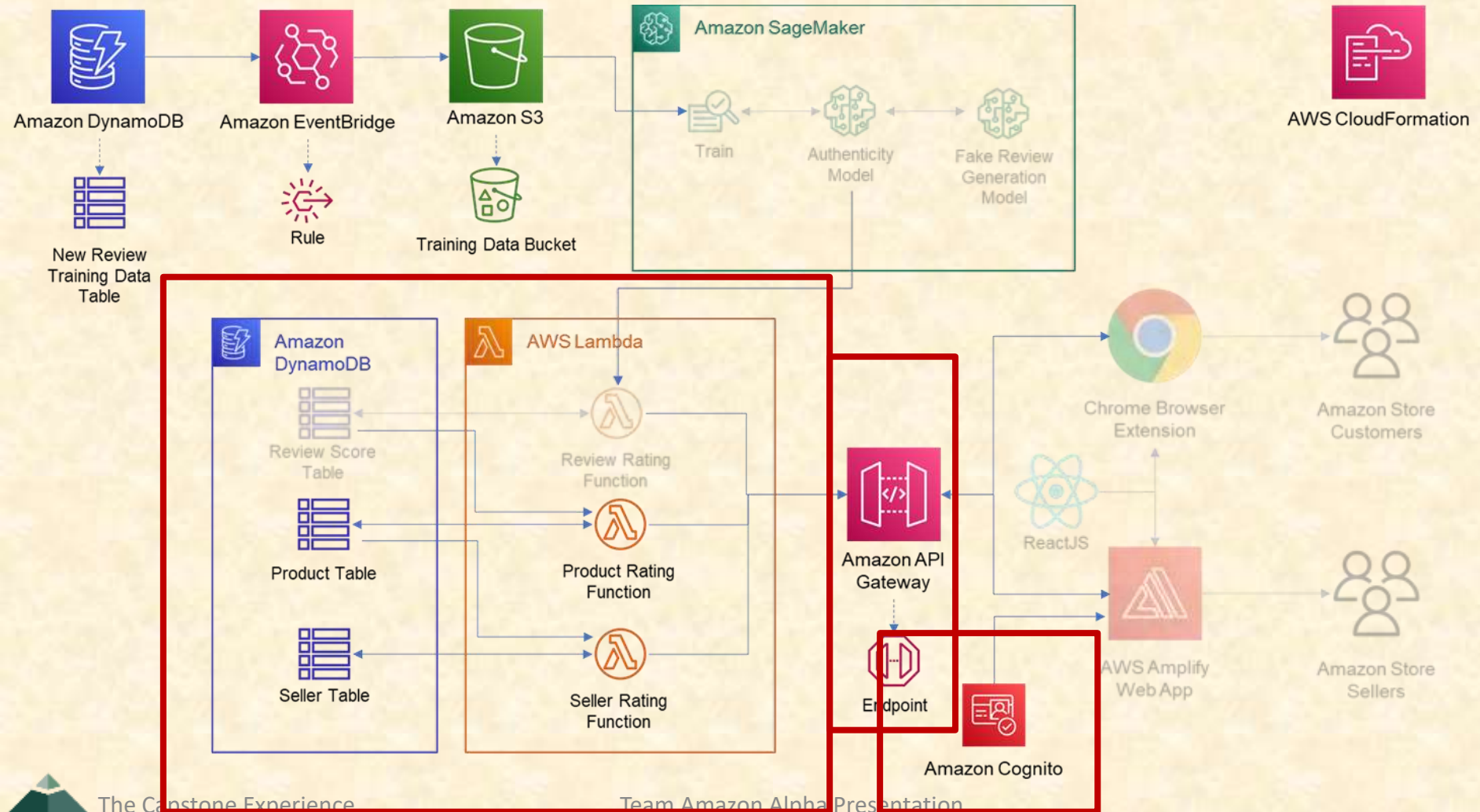
Rated 5 out of 5 by JayC from Great deal I am very impressed with the quality and quality of the Canon EOS 35mm f/1.2 E DSLR. You can also use it for the Sony STIF. I have a 35mm f/1.2 from the Sony STIF but no 50mm f/1.1 E lens. This camera is also better than other high end lenses which I own. I would buy again for that to work and use for other things or have it back.



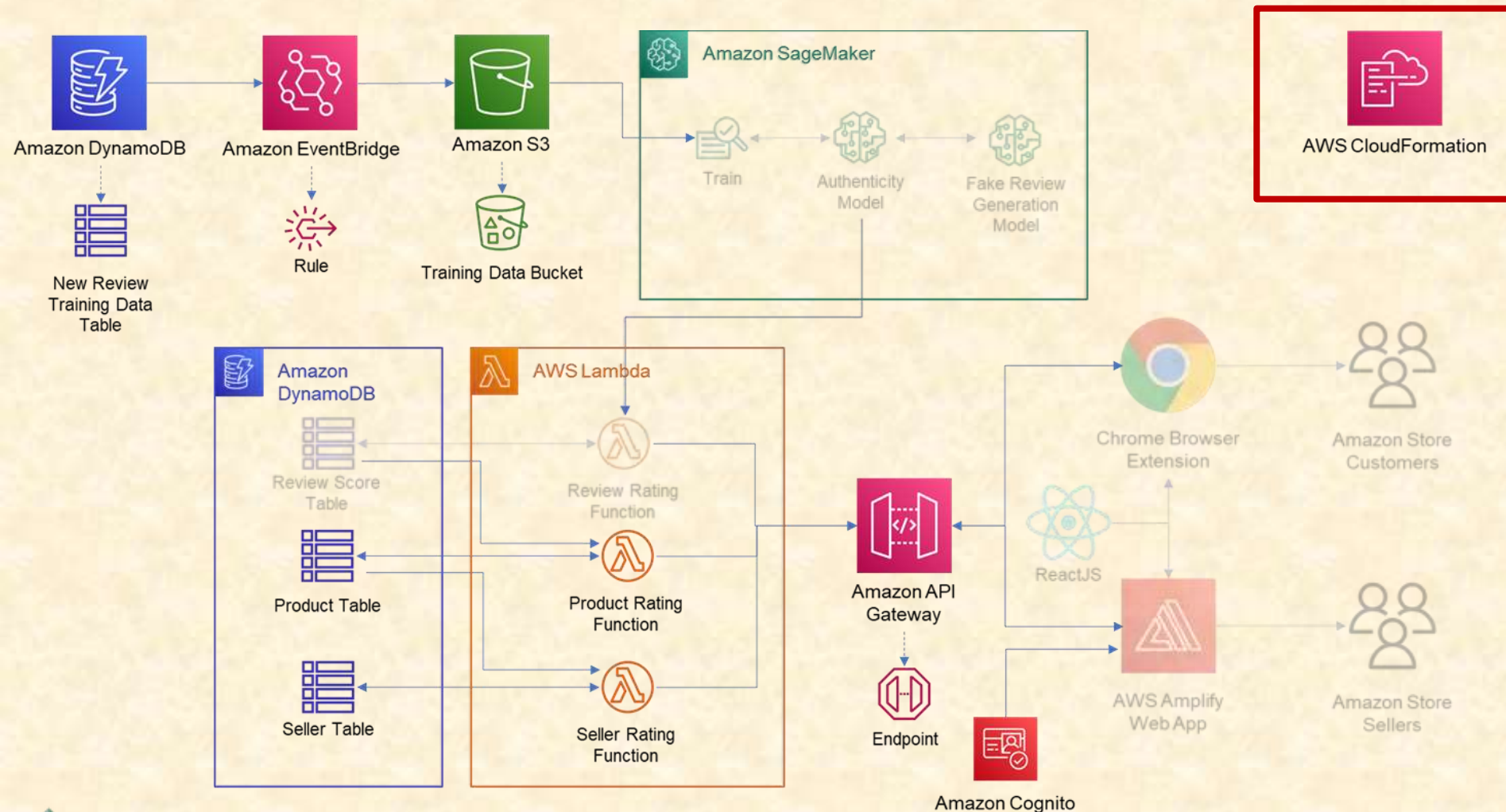
What's Left: Current Status



What's Left: Beta



What's Left: Deliverable



Questions?

?

?

?

?

?

?

?

?

?



End of slide show, click to exit.

ML Preprocessing Explained

- `SkLearn.CountVectorizer(analyzer)`
 - Tokenize text features
 - Analyzer picks out stopwords, punctuation, etc
- Term Frequency – Inverse Document Frequency
 - Word “Originality”
 - Tf: times word **t** appears in doc **d**
 - Idf: # of docs word appears in

Scikit-Learn

- $\text{IDF}(t) = \log \frac{1+n}{1+\text{df}(t)} + 1$

Standard notation

- $\text{IDF}(t) = \log \frac{n}{\text{df}(t)}$

In implementation, must account for
divide-by-zero errors

