

# Machine Learning @ Amazon

Rajeev Rastogi

Director, Machine Learning



# Key Takeaways

- ML is enabling smart ecommerce
  - Product recommendations, demand forecasting, search, classification, matching, ...
- Learning semantically rich representations critical for many current tasks and future AI applications
  - Online advertising, product search, **question answering**, **product recommendations**, fake reviews detection
  - Conversational systems, video metadata generation, content summarization
- Several techniques for learning representations
  - **Deep Learning**, **Probabilistic Graphical Models**, Tensor Decomposition
  - Leverage diverse signals/data

# Outline

- Applications of ML @ Amazon
- Question answering
- Product size recommendations

# Amazon Marketplace

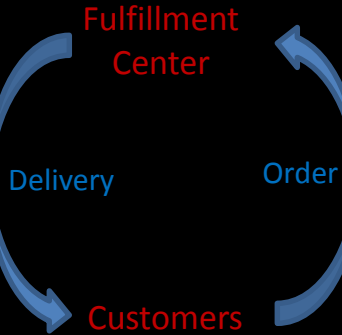


Ship  
Inventory



Sellers

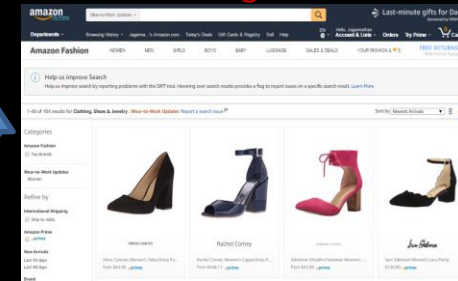
List  
products



Catalog



Browse, Search &  
Recommendations



# Using ML to Address Top Business Priorities

- Increasing product selection
- Lowering prices
- Reducing delivery times
- Eliminating friction
- Maintaining customer trust

# Numerous ML Applications

- Demand Forecasting
- Vendor Lead Time Prediction
- Pricing
- Packaging
- Substitute Prediction

Retail



- Product Recommendation
- Product Search
- Product Ads
- Customer Targeting

Customers



- Seller Nudges
- Fraud Detection
- Predictive Help
- Seller Lead Generation

Sellers



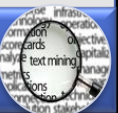
- Product Classification
- Product matching
- Reviews Ranking and Analysis
- Fraudulent Reviews detection

Catalog



- Named-Entity Extraction/Xray
- Summarization
- In-Book Search
- Plagiarism Detection

Text



- Visual Search
- Product Image Quality
- Brand Tracking

Images



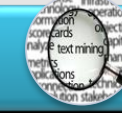
- Predicting Resource Requirements
- DDoS detection
- Reputation computation of Mturk workers

AWS



- Automatic speech recognition
- Natural Language Understanding
- Dialog Management

Speech



# Product Demand Forecasting

## Problem

- *Given past sales of a product in every region of the world, predict regional demand up to one year into the future*



# Product Demand Forecasting

## Problem

- *Given past sales of a product in every region of the world, predict regional demand up to one year into the future*

## Challenges

- **Scale:** Hundreds of millions of products in Amazon catalog!
- **New products:** No past demand!
- **Sparsity:** Huge skew – many products sell very few items
- **Seasonal:** Demand for some products exhibits seasonal patterns
- **Demand spikes:** Huge variation due to external events
- **Distributions:** Future is uncertain → predictions must be distributions



# Product Search

## Problem

- Given a partial user query, find the relevant products to display to user in search results

The screenshot shows the Amazon India search results page for the query "machine learning". The page displays 1-50 of 1,413 results. The search bar at the top shows the query "machine learning" and a "Go" button. The page is sorted by "Relevance".

**Wiley Big Data Sale**  
Tame the big data tidal wave with Wiley guides to data visualization, analytics, data mining, and more.

**Related Searches:** artificial intelligence, deep learning, neural networks.

**Refine by:**  
New Releases  
Last 30 days (10)  
Last 90 days (10)  
Coming Soon (3)  
Book Series  
Author  
Kindle Unlimited

**Product Grid:**

Product Title	Price	Kindle Edition	Rating
Machine Learning: The Art and Science of Algorithms that Make Sense of Data (Sep 2012)	Rs. 2,295.99	Kindle Edition	★★★★★ (15)
Machine Learning: Hands-On for Developers and Technical Professionals (Oct 20, 2014)	Rs. 1,882.34	Kindle Edition	★★★★★ (4)
Machine Learning: A Probabilistic Perspective (Adaptive Computation and Machine Learning series) (Sep 7, 2012)	Rs. 4,388.10	Kindle Edition	★★★★★ (1) (2)
Machine Learning: An Algorithmic Perspective, Second Edition (Chapman & Hall/CRC Machine Learning & Pattern Recognition... (Oct 14, 2014)	Rs. 3,726.85	Kindle Edition	★★★★★ (2)
Practical Machine Learning: Innovations in Recommendation (Apr 17, 2014)	Rs. 6.00	Kindle Edition	★★★★★ (1) (6)
Machine Learning with R (Oct 25, 2013)	Rs. 388.40	Kindle Edition	★★★★★ (1) (3)
Understanding Machine Learning: From Theory to Algorithms (May 14, 2014)	Rs. 1,539.74	Kindle Edition	★★★★★ (1) (7)
DATA MINING: An Introduction to Statistical Learning			
Machine Learning			
Machine Learning with Spark			
Machine Learning			
Foundations of Machine Learning			
Big Data Now			

# Product Search

## Problem

- *Given a partial user query, find the relevant products to display to user in search results*

## Challenges

- **Scale:** Hundreds of millions of products in Amazon catalog
- **Real-Time Prediction:** Search requires low-latency (<40ms)
- **Query analysis:** Understand semantics, identify phrases, classify into product category (e.g. red apple iphone vs red apples)
- **Intent detection:** Is customer researching product or looking to buy product
- **Knowledge Graph:** Structured entity data + related products in response to search queries

# Product Classification

## Problem

- Given a product description from a seller, map it to the appropriate leaf node in product taxonomy

amazon  
Jaganna...'s Amazon.com Today's Deals Gift Cards Sell Help

Valentine's Day Gifts Shop now

Shop by Department - Search Clothing, Shoes & Jewelry - Go Hello, Jagannathan Your Account Try Prime Cart - Wish List -

Amazon Fashion Women - Men - Girls - Boys - Baby - Luggage - Sales And Deals - Your #s - FREE RETURNS

Clothing, Shoes & Jewelry > Girls > Shoes > Outdoor > Rain Boots

Polo Ralph Lauren  
**Polo by Ralph Lauren Ralph Rain Boot (Toddler/Little Kid/Big Kid)**  
★★★★☆ - 27 customer reviews

Price: \$49.90 - \$65.00 & FREE Returns on some sizes and colors. Details

Size:  
Select Size Chart | Fit: As expected (62%)

Color: Pink/Green

- Rubber
- Rubber sole
- Shaft measures approximately 11" from arch
- Heel measures approximately 1"
- Boot opening measures approximately 13" around
- Waterproof
- Logo detailing
- This rain boot from Polo by Ralph Lauren features a charming buckle detail and a grippy tread.

To buy, select Size  
Choose from options to the left

Add to Cart

Add to Wish List

# Product Classification

## Problem

- *Given a product description from a seller, map it to the appropriate leaf node in product taxonomy*

## Challenges

- **Scale:** Hundreds of millions of products and thousands of classes in product taxonomy
- **Products vs Accessories:** Hard to distinguish between products and its accessories (e.g. laptop vs laptop battery)
- **Incorrect/Missing data:** Some attribute values may be wrong, missing or inadequate (e.g. short titles)
- **Training data quality:** Examples may be wrongly labeled, some classes may have very few examples

# Product Matching

## Problem

- Given product information (title, description, price, etc.), find duplicate product listings in Amazon catalog

The screenshot shows an Amazon search results page for the product "Ninja Master Prep QB900B". The page displays 14 results, with the first three listings highlighted by red rounded rectangles to illustrate the problem of duplicate listings. The first listing is "Ninja Master Prep (QB900B)" priced at \$11.20. The second listing is "Ninja QB900B Master Prep™ Food & Drink Mixer" priced at \$59.99. The third listing is "Euro-Pro Sales QB900B Ninja Master Food Prep Blender/Processor - Quantity 2" priced at \$35.70. The page also includes a left sidebar with filters for category, price, and customer reviews, and a top navigation bar with the Amazon logo and search bar.

amazon  
Jaganna...s Amazon.com Today's Deals Gift Cards Sell Help

Valentine's Day Gifts

Stop by Department - Search Appliances - Ninja Master Prep QB900B Go Hello, Jagannathan Your Account - Try Prime - Cart -

Appliances Best Sellers Refrigeration Cooking Washers & Dryers Dishwashers Parts & Accessories

14 results for Appliances: "Ninja Master Prep QB900B"

Show results for

Any Category

Appliances

Blenders (3)

Food Processors (7)

Countertop Blenders (2)

See more

Refine by

Eligible for Free Shipping

Free Shipping by Amazon

Avg. Customer Review

5 stars & up (3)

4 stars & up (3)

3 stars & up (3)

2 stars & up (3)

1 star & up (3)

New Arrivals

Last 30 days (3)

Last 90 days (5)

Price

\$25 to \$50 (7)

\$50 to \$100 (6)

\$100 to \$200 (1)

\$ to \$

Discount

10% Off or More (6)

25% Off or More (4)

50% Off or More

70% Off or More

See Style Options

Ninja Master Prep (QB900B)

\$11.20 ~~\$46.00~~ Prime

Get it by Monday, Feb 9

More Buying Choices

\$31.20 new (46 offers)

\$26.52 used (23 offers)

FREE Shipping on orders over \$35

Best Seller in Food Processors

5 stars & up - 3,300

Ninja QB900B Master Prep™ Food & Drink Mixer

\$59.99

More Buying Choices

\$43.99 used (2 offers)

FREE Shipping

Euro-Pro Sales QB900B Ninja Master Food Prep Blender/Processor - Quantity 2

\$35.70

More Buying Choices

\$72.15 new (3 offers)

FREE Shipping

Ninja Master Prep Food Processor and Drink Mixer in Gray

\$51.30

Only 10 left in stock - order soon.

Master Prep (QB900B) 9Z7102XL

669.00 \$53.99 + \$6.49 shipping

Ninja Master Prep Pulsating Food and Drink Maker

\$66.97 new (2 offers)

# Product Matching

## Problem

- *Given product information (title, description, price, etc.), find duplicate product listings in Amazon catalog*

## Challenges

- **High-Precision requirement:** Incorrect matching leads to poor user experience
- **Variations:** Some variations (e.g., color) are insignificant while others (e.g., movie sequels) are not
- **Diverse formats:** Attribute values for the same product may be represented differently
- **Incorrect/Missing Data:** Some attribute values may be missing or wrong in product descriptions

# Information Extraction from Reviews

## Problem

- Generate review summaries, extract product attribute ratings from reviews

The screenshot displays the Amazon product page for the Panasonic DMC-GH4EB-K Compact System Camera. The page is divided into several sections:

- Customer Reviews:** Shows a 4.7 out of 5 stars rating based on 21 reviews. A bar chart indicates the distribution of star ratings: 5 stars (17), 4 stars (2), 3 stars (1), 2 stars (1), and 1 star (0).
- Top Customer Reviews:** Features a review by Gozaradio (29 Aug. 2015) with a 5-star rating, titled "The gold standard in budget".
- Review summary from DPReview:** Includes a quote from Richard Butler, Rishi Sanyal, Jeff Keller, and EOSHD: "Despite the increased competition, the GH4 remains the king of accessible stills/movie hybrids. The enhanced movie capabilities are well integrated such that both types of shooting work well, without impeding the other. The result is a remarkably capable all-in-one package".
- Conclusion:** A circular progress indicator shows 85% with the text: "Despite increased competition, the GH4 remains the most complete stills/video camera on the market. It's a competitive stills camera in its own right but the attention to detail when it comes to the provision of movie-shooting tools and capabilities makes it a stand-out camera."
- Reasons to buy:**
  - Solid, comfortable-to-use camera body
  - Excellent still image quality
  - Industry-leading video capabilities
  - Highly detailed video capture
  - Good viewfinder and rear screen
- Things to consider:**
  - Focus tracking is poor at subject identification
  - Focus peaking is often too subtle to assess focus point
  - No Auto ISO in manual exposure video shooting mode
  - Wi-Fi feature is a little awkward to set up
- Suggested for:** Enthusiast film makers and anyone who cares as much about their movie footage as their still images.
- Not suggested for:** Sports photographers or landscape shooters looking for the pinnacle of image quality.
- Attribute Rating Table:** A table comparing the camera's performance across various attributes on a scale from POOR to EXCELLENT.

Attribute	Rating
Build quality	High
Ergonomics & handling	High
Features	High
Metering & focus accuracy	High
Image quality (RAW)	High
Image quality (JPEG)	High
Low light / high ISO performance	High
Viewfinder / screen rating	High
Performance	High
Movie / video mode	High
Connectivity	High
Value	High

# Information Extraction from Reviews

## Problem

- *Generate review summaries, extract product attribute ratings from reviews*

## Challenges

- **Diverse attributes:** Product attributes may vary across different categories
- **Synonyms:** Different terms may refer to same attribute (sound, audio)
- **Informal style:** User reviews have an informal linguistic style
- **Stylistic variations:** Linguistic style varies between users
- **Sentiment analysis:** Gauging sentiment may require deep parsing of sentences



# Product Recommendations

## Problem

- Discover products by recommending the right product, to the right customer, in the right place, at the right time

The screenshot shows the Amazon product page for 'Lady Luck: The Theory of Probability' by Warren Weaver. The page includes a search bar, navigation links, and a 'Look inside' feature. A red box highlights the 'Your Recently Viewed Items and Featured Recommendations' section at the bottom, which displays a carousel of related books:

- Amazon Fire TV User Guide (2014 Edition)
- Mathematics for the... (Morris Kline)
- How To Learn And Memorize Math... (Anthony Mazer)
- Calculus: An Intuitive and... (Morris Kline)
- Fourier Series (George P. Tolstov)
- ALGEBRA IN WORDS: A Guide of... (Gregory Bullock)
- Physics in Quantities and Examples (Metin Baltas)
- THE ART OF... (Jessica Riley)

# Drones

## Problem

- *Safely deliver packages to customers homes within 30 minutes using drones*



# Robotics

## Problem

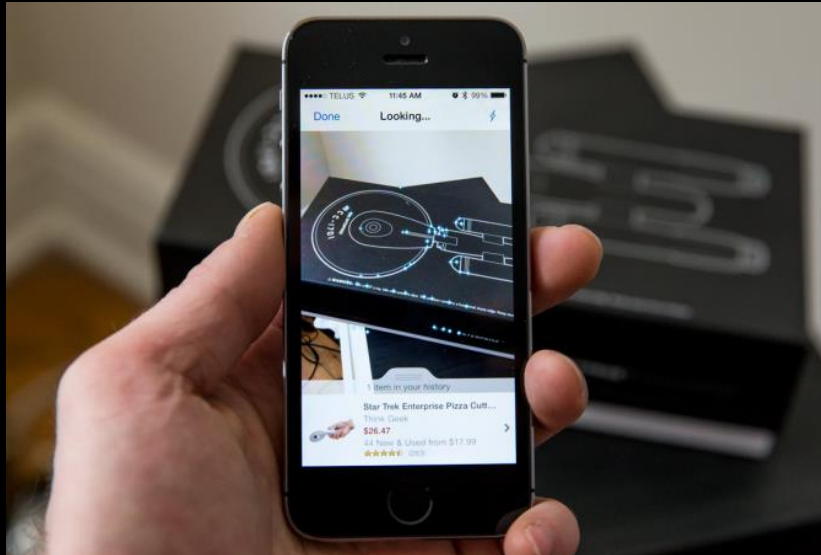
- Automate picking, stowing and transport of products in Amazon Fulfillment Centers



# Visual Search

## Problem

- Retrieve images from the Amazon catalog that are visually similar to a given product



# Voice Recognition

## Problem

- *Provide a voice interface to shop for products, perform tasks, answer questions, and carry out conversations*

## The Amazon Echo family



# Outline

- Applications of ML @ Amazon
- Question answering
- Product size recommendations

# Amazon Product Pages

- Amazon product pages contain a wealth of information

The image displays three sequential screenshots of an Amazon mobile product page for a GE A730 camera. The first screenshot shows the product image, a 4.8-star rating with 95 reviews, and a list of features including 7MP digital camera, 3X optical zoom, and image stabilization. The second screenshot shows a 'Product Information Black' table with details like dimensions (5.6 x 4.1 x 2.6 inches), weight (1.1 pounds), and manufacturer (General Imaging). The third screenshot shows a 'Customer Questions' section with several Q&A entries regarding video recording, interior photography, and memory card compatibility.

**Product Information Black**

Product Dimensions	5.6 x 4.1 x 2.6 inches
Item Weight	1.1 pounds
Shipping Weight	1.2 pounds
Manufacturer	General Imaging
ASIN	B000QGBCVY

**Customer Questions**

Q: **does it record videos well?**  
A: It does shoot video, kind of well, the lighting is very bad on video I'd say, and for it...  
dianna decoste | 3 years ago

Q: **I a looking for a camera that takes good interior pictures, wide angle for realtor photos,...**  
A: Sure it's a great little camera, I love it indoors and out.  
Lois S. | 9 months ago

Q: **Will this use a 32gb sdhc memory card?**  
A: Yep  
albert | 2 years ago

Q: **Does it have a timer?**  
A: It has a three-mode timer: 2 second, 10 second and 15 second. Smile mode...

# Question & Answering Bot

- Question answering interface to make it easy for users to find information on product page

The image displays three sequential screenshots of an Amazon mobile application interface, illustrating a question-answering bot's integration into a product page.

**Left Screenshot: Product Page**  
The top navigation bar shows the Amazon logo and search icons. The main content area features a product image of a GE A730 7MP Digital Camera. Below the image, the product title is "GE A730 7MP Digital Camera with 3x Optical Zoom (Black) (OLD MODEL)" with a 4.5-star rating and 95 reviews. A red-bordered box highlights a "Ask us about this item" button with the subtext "Enter a question, and we'll search this page".

**Middle Screenshot: Product Information**  
The "About this item" section is visible, showing a list of features and a "Product Information Black" table. The table contains the following data:

Product Information	Value
Product Dimensions	5.6 x 4.1 x 2.6 inches
Item Weight	1.1 pounds
Shipping Weight	1.2 pounds
Manufacturer	General Imaging
ASIN	B000QGBCVY

**Right Screenshot: Customer Questions**  
The "Customer Questions" section is shown, featuring a search bar and a list of questions and answers. The questions include:

- Q: does it record videos well?**  
A: It does shoot video, kind of well, the lighting is very bad on video I'd say, and for it...  
dianna decoste | 3 years ago
- Q: I a looking for a camera that takes good interior pictures, wide angle for realtor photos,...**  
A: Sure it's a great little camera, I love it indoors and out.  
Lois S. | 9 months ago
- Q: Will this use a 32gb sdhc memory card?**  
A: Yep  
albert | 2 years ago
- Q: Does it have a timer?**  
A: It has a three-mode timer: 2 second, 10 second, and 15 second. Smile mode...



# Product Feature Questions

The image displays three sequential screenshots from an Amazon mobile application, illustrating how product features are highlighted in customer reviews. The first screenshot shows the product page for a Canon PowerShot SX720 HS camera, with a red box highlighting the 'Item Weight: 8.6 ounces' in the product information table. The second screenshot shows a customer review for a DeeXop HD Mini Digital Camera, with a red box highlighting the text 'For the price this camera is good value for money for an amateur protographer.' The third screenshot shows another review for the same DeeXop camera, with a red box highlighting the text 'Brought as a birthday present for my mum. She loves how easy it is to use, compact and good photo quality. For the price this camera is good value for money for an amateur protographer.' Each screenshot includes a search bar, product image, and navigation buttons.

**Product Information Table:**

Product information	
Product Dimensions	2.5 x 1.4 x 4.3 inches
Item Weight	8.6 ounces
Shipping Weight	1.2 pounds
Manufacturer	Canon
ASIN	B01BV14140
Origin	Japan
Manufacturer reference	1070C001
Batteries	1 Lithium ion batteries

**Customer Reviews:**

- Review 1:** Brought as a birthday present for my mum. She loves how easy it is to use, compact and good photo quality. For the price this camera is good value for money for an amateur protographer.
- Review 2:** Still a great buy. My son loves it. Excited Great camcorder! Great package of accessories. Shipped fast. Great price! Pleased with the company, will order from them again.
- Review 3:** Wrks great.

# Product Comparison/Compatibility Questions

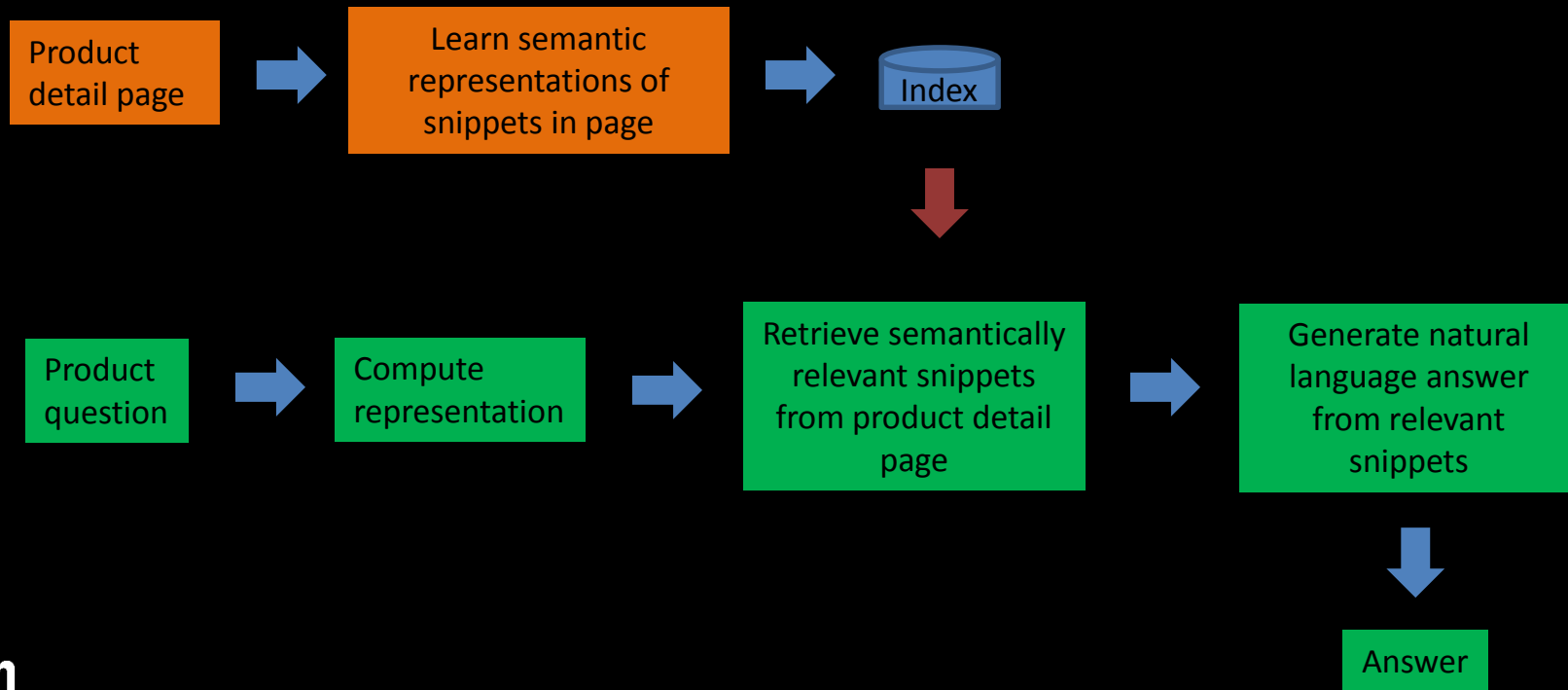
The image displays four screenshots of Amazon product pages, each showing a customer question and an answer regarding camera compatibility. The screenshots are arranged in a 2x2 grid, with a dashed purple border around the entire set.

- Top Left Screenshot:** Product page for "Canon EOS Rebel T5i EF-S 18-55 IS STM Kit". A question asks, "What is the difference between this camera and the Rebel SL1?". The answer states: "The SL1 is more compact and the specs are not as good as the T5i. However, it's still a good camera if you are looking to save some money." This answer is highlighted with a red box.
- Top Right Screenshot:** Product page for "Canon EOS Rebel T6 Digital SLR Camera Kit with EF-S 18-55mm and EF 75-300mm Zoom Lenses (Black)". A question asks, "Is it compatible with canon lenses?". The answer states: "The EOS Rebel T6 camera is compatible with all Canon lenses in the EF and EF-S lineup." This answer is highlighted with a red box.
- Bottom Left Screenshot:** Product page for "Canon EOS Rebel T5i EF-S 18-55 IS STM Kit". A question asks, "How does it compare with rebel SL1?". The answer states: "The SL1 is more compact and the specs are not as good as the T5i." This answer is highlighted with a red box.
- Bottom Right Screenshot:** Product page for "Canon EOS Rebel T6 Digital SLR Camera Kit with EF-S 18-55mm and EF 75-300mm Zoom Lenses (Black)". A question asks, "Compatible with the full line of Canon EF and EF-S lenses?". The answer states: "The EOS Rebel T6 camera is compatible with all Canon lenses in the EF and EF-S lineup." This answer is highlighted with a red box.

# Key Challenges

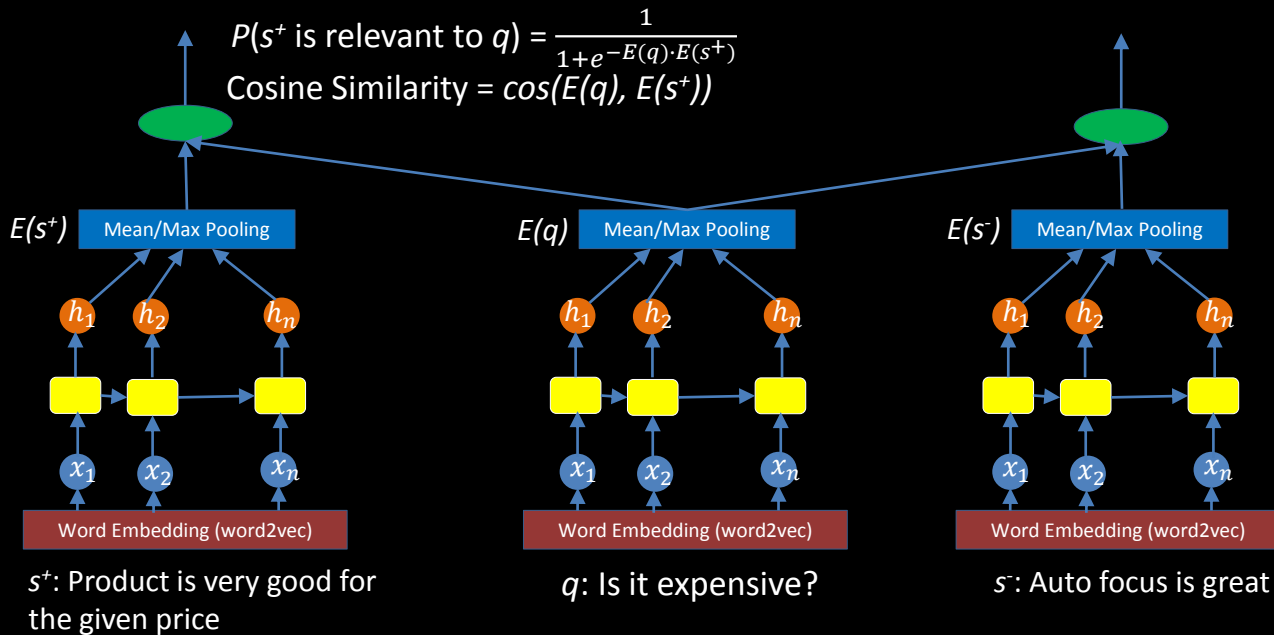
- Question understanding
  - “What is ISO?” vs “What is the ISO [of this camera]?”
- Semantic matching
  - “cost”, “price”, “bang for buck”, “expensive”, “cheap”
- Natural language answer generation
  - e.g. “This is great value for money” for question “Is this expensive?”
- High precision (>90%) requirement
- Data availability
  - “Will this suitcase fit in the overhead of an airplane?”
- Data quality
  - “Dimensions: 1x1x1 inches” for the question “How big is it” on a “Tripod page”

# Question Answering System Architecture



# Learning Semantically Rich Representations

- Training examples: <question ( $q$ ), relevant snippet ( $s^+$ ), irrelevant snippet ( $s^-$ )> triples
- **Triplet network**



# Results for Different Loss Functions

- Learn question and snippet representations to minimize the following loss functions:
  - Log loss [EMNLP 2015]
$$-\log \frac{1}{1 + e^{-E(q) \cdot E(s^+)}} - \log \frac{1}{1 + e^{E(q) \cdot E(s^-)}}$$
  - Siamese loss [Wang et al. 2014]
$$\max\{0, M - (\cos(E(q), E(s^+)) - \cos(E(q), E(s^-)))\}$$
  - Twin loss
$$\max\{0, M_1 - \cos(E(q), E(s^+))\} + \max\{0, \cos(E(q), E(s^-)) - M_2\}$$
- Metric: Precision at rank 1 (P@r1)
- Results:

Loss function	Baseline	Log loss	Siamese loss	Twin loss
P@r1	56.8%	84.6%	96.1%	97.04%

# Qualitative Results

Question	Matching Snippet
Is this camera good for pictures at a basketball game?	Works great for <b>sports</b> photography
What is the price ?	This item <b>costs</b> \$100.00. To see tax and shipping, add to cart
How big is it?	Item <b>dimensions</b> : 3 x 3.28 x 4.37 inches
How good is stabilization?	EVERY image came out <b>blurry</b> (and I held the camera still in a well-lit room).
Will it fit on Olympus air?	Fits very well the Olympus Air OA-01
How much weight can it hold?	Item <b>weight</b> : 2.2 pounds
What is the color of the paper on which the photo is printed?	<b>the color of the camera and the pictures are great.</b>

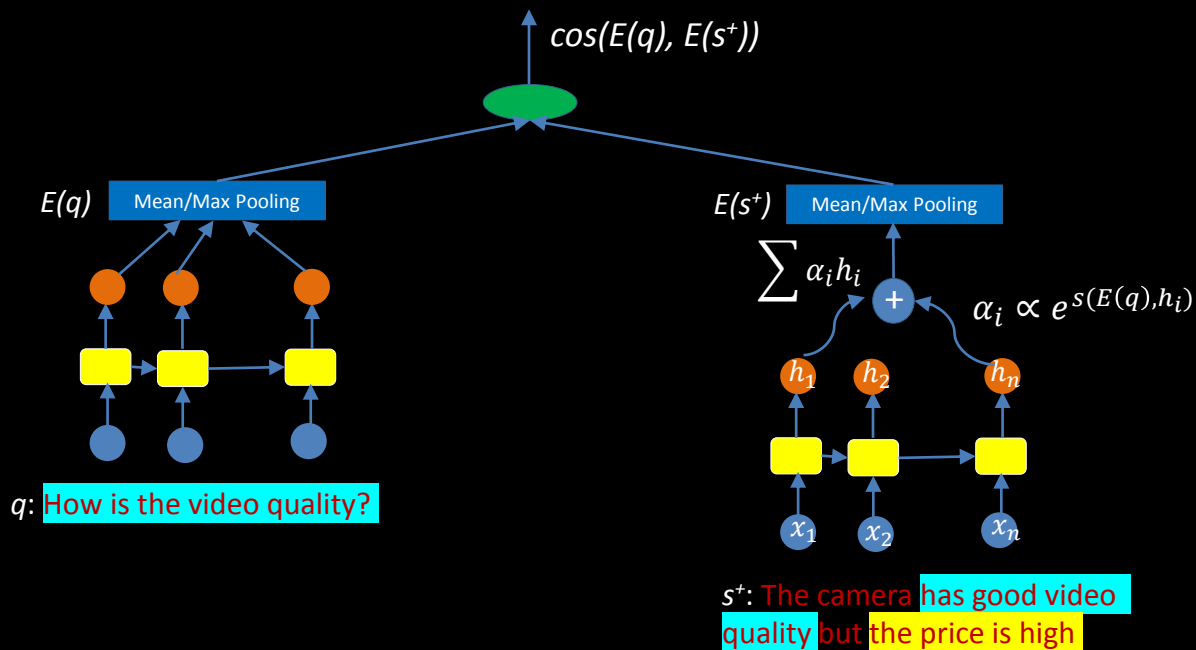
# Training Dataset Generation

- Raw Data
  - Corpus of Amazon customer reviews, product descriptions, community QA
  - Total size is 147.6M records comprising 3.68M word tokens
  - Trained word2vec embeddings of 200 dimensions
- Training examples: triplets  $(q, s^+, s^-)$ 
  - Attribute selection: mined attributes using term frequency (~130 for cameras)
    - price, image stabilization, zoom, resolution, etc.
  - Question templates  $(q)$  for attributes
    - price : “what’s the price”, “what is the cost”, etc.
  - Relevant snippets  $(s^+)$ : defined set of matching words
    - price: “price”, “expensive”, “cheap”, “worth”, “money”, “pricey”, “investment”, “buck”, “cheaply”
  - Irrelevant snippets  $(s^-)$ : used negative sampling
  - Over 10M triplets



# Learning Representations with Attention

- Only consider relevant portions of snippets when learning representations [Bahdanau et al. 2015]



# Highlighting Words with High Attention Weights

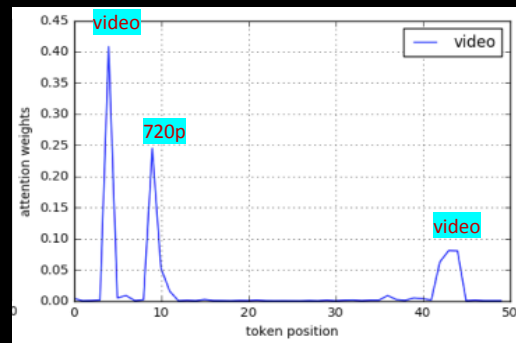
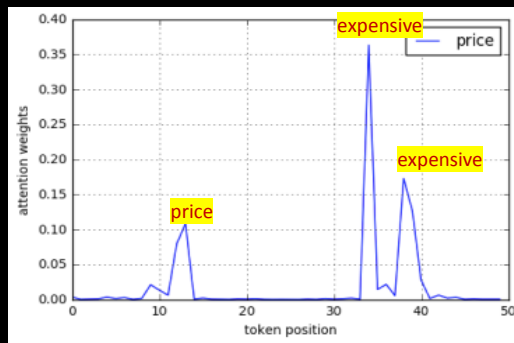
## Review Statement:

The package is good **video** quality is good for **720p** but the **price** is excellent for what you get - especially if you do not want all the whistles and bells of the more **expensive** gopro 2-4x more **expensive** and the quality of **video** is superb and great.

## Questions:

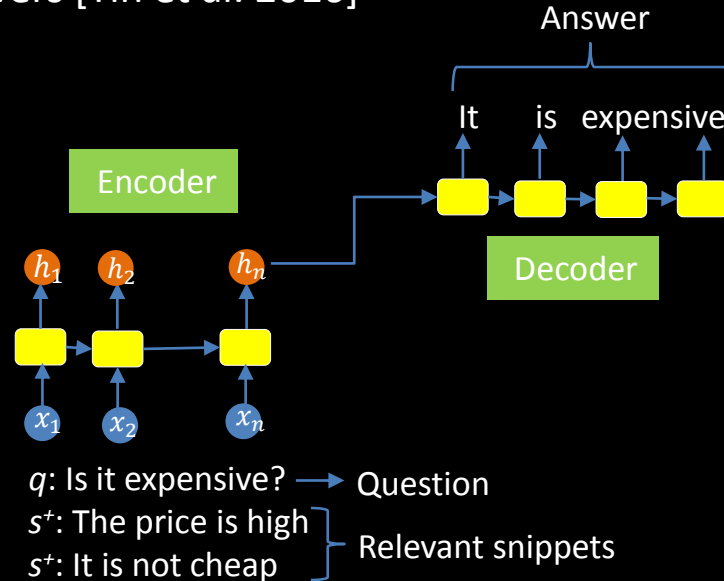
Question1: **What is the price?**

Question2: **How good is the video?**



# Generating Natural Language Answers

- Use sequence-to-sequence encoder-decoder model to generate final answers [Yin et al. 2016]



# Outline

- Applications of ML @ Amazon
- Question answering
- Product size recommendations

# Size Recommendation Problem

- Given a customer and product, recommend product size that would best fit the customer

Helen's Pinkmartini  
Helen's Men's Pinkmartini 7 Colors Light Up High Top Sports Sneakers  
★★★★☆ 269 customer reviews | 269 answered questions  
Price: \$26.99 - \$49.99  
Fit: As expected (79%)  
Size:  
Select Size Chart  
4 D(M) US  
4.5 D(M) US  
5.5 D(M) US  
6 D(M) US  
7 D(M) US  
7.5 D(M) US  
8.5 D(M) US  
9 D(M) US  
10 D(M) US  
10.5 D(M) US  
11 D(M) US  
12 D(M) US  
35 EU  
36 EU  
37 EU  
38 EU  
39  
40 EU  
41 EU  
42 EU  
43 EU  
44 EU  
45 EU  
weaker featuring four-eyelid face-up closure for easy on and off. Stretch  
comfort, vent holes for breathability  
durable rubber sole against abrasion  
fit and comfort, stability and speed  
product information.  
viewed this item also bought  
Allroad \$27.99 - \$33.99  
JustCrat \$28.00 - \$68.99  
ITURBOS \$39.99

# Motivation

- No standardization of sizes across brands and locales for product categories such as shoes and apparel
- This leads to users making incorrect purchases, and then returning products
- Products belonging to shoe and apparel categories have high return rates due to fit issues
- Example:
  - Reebok size mapping convention: 6 = 15cm, 7 = 17cm, 8 = 21cm
  - Nike size mapping convention: 6 = 16cm, 7 = 18cm, 8 = 22cm

# Key Challenges

- Scale: hundreds of millions of customers and products
- Data sparsity: bulk of users/products have very few purchases
- Cold start scenarios: new customers/products
  - User features: demographics (age, gender), location
  - Product features: catalog size, title, brand, product type
- Multiple personas: each customer  $i$  may involve multiple personas
  - E.g., family members sharing an account
  - Personas may have widely varying sizes

# Our Approach

- Learn true (latent) size for each customer, product
  - True size for customer corresponds to the physical size of the customer (for shoes, it would be the feet size)
  - True size for product corresponds to its physical size
- Leverage past customer transactions  $T = \{(i, j, y_{ij})\}$ 
  - $y_{ij}$  takes ordinal values {small, fit, large}

	Adidas (9)	Nike (8)	Reebok (8)	Nike (9)
Customer 1	large		fit	?
Customer 2		small		fit
Customer 3	fit		small	?
Customer 4		fit		large

→ Catalog size

→ Predict fit outcome



# Our Approach (Contd)

- Notation
  - Latent size for customer  $i$ :  $s_i$
  - Latent size for product  $j$ :  $t_j$
  - Catalog size for product  $j$ :  $c_j$
- Model likelihood of fit as a function of the difference between customer and product latent sizes

$$P(y_{ij} = fit) \propto f(s_i - t_j)$$

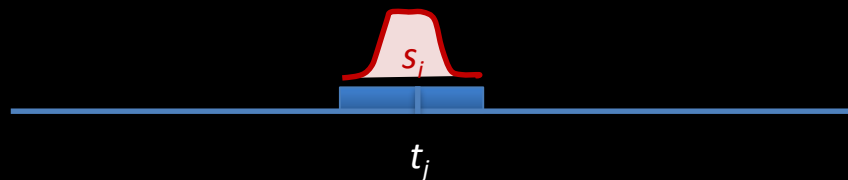
- Recommend product  $j$  with highest fit likelihood  $P(y_{ij} = fit)$  to customer  $i$

# Bayesian Modeling Benefits

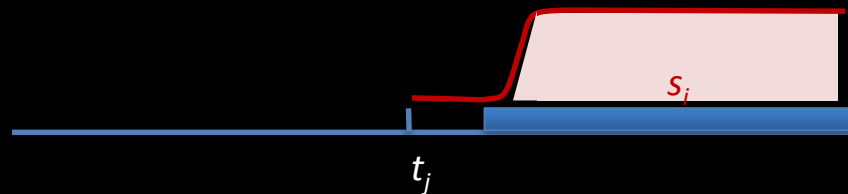
- Handles data sparsity by placing priors on latent size variables
- Models uncertainty in inferred latent sizes
  - Estimates posterior distribution of latent size variables
  - Fit probability is obtained by averaging over posterior size distribution
- Model can capture all the available data
  - Observations: transaction outcomes, customer and product features
  - Hidden variables: latent sizes, customer personas
- Efficient techniques for approximating posterior distributions of latent size variables

# Intuition

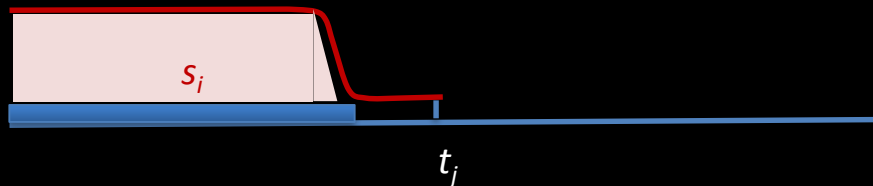
- Transaction  $(i, j, fit)$   $\rightarrow$  very likely that  $s_i$  and  $t_j$  are close



- Transaction  $(i, j, small)$   $\rightarrow$  very likely that  $s_i$  is much larger than  $t_j$



- Transaction  $(i, j, large)$   $\rightarrow$  very likely that  $s_i$  is much smaller than  $t_j$



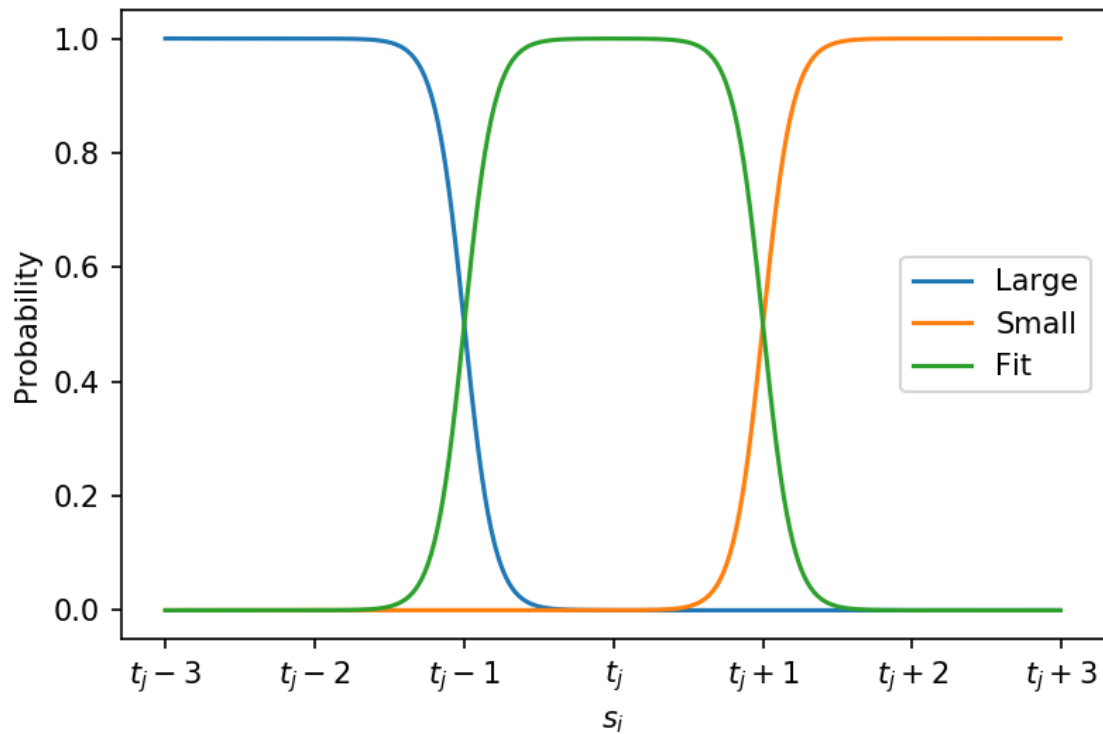
# Data Likelihood

$$P(y_{ij} = \textit{small} | s_i, t_j) = \frac{1}{1 + e^{-\alpha(s_i - t_j) + b_1}}$$

$$P(y_{ij} = \textit{fit} | s_i, t_j) = \frac{1}{1 + e^{\alpha(s_i - t_j) - b_1}} \cdot \frac{1}{1 + e^{-\alpha(s_i - t_j) + b_2}}$$

$$P(y_{ij} = \textit{large} | s_i, t_j) = \frac{1}{1 + e^{\alpha(s_i - t_j) - b_1}} \cdot \frac{1}{1 + e^{\alpha(s_i - t_j) - b_2}}$$

# Data Likelihood



# Generative Model

for each customer  $i$ ,

draw latent size  $s_i \sim N(\mu_s, \sigma_s^2)$

for each product  $j$ ,

draw latent size  $t_j \sim N(c_j, \sigma_t^2)$

for each transaction  $(i, j, y_{ij}) \in T$ ,

select  $y_{ij} = \textit{small}$  with probability  $P(y_{ij} = \textit{small} | .)$

select  $y_{ij} = \textit{fit}$  with probability  $P(y_{ij} = \textit{fit} | .)$

select  $y_{ij} = \textit{large}$  with probability  $P(y_{ij} = \textit{large} | .)$

# Bayesian Inference

- Let  $\beta$  be the vector of latent sizes

$$\beta = (s_1, \dots, s_c, t_1, \dots, t_p, 1)^T$$

- Posterior distribution

$$\begin{aligned} P(\beta|T) &\propto P(T|\beta) \cdot P(\beta) \\ &\propto \prod_{(x,y) \in \mathcal{D}} \frac{e^{y\beta^T \cdot x}}{1 + e^{\beta^T \cdot x}} \cdot \prod_i N(s_i | \mu_i, \sigma_s) \cdot \prod_j N(t_j | c_j, \sigma_t) \end{aligned}$$

$0/1$                        $(0, \dots, 0, \alpha, 0, \dots, 0, -\alpha, 0, \dots, 0, -b_1/b_2)^T$

- Not available in closed form due to logistic likelihood terms and Normal priors

# Polya-Gamma Augmentation [Polson et al. 2013]

- Introduce Polya-Gamma latent variable  $w \sim PG(0,1)$  for every  $(x, y) \in \mathcal{D}$
- Define the joint likelihood distribution

$$P(w, y|x, \beta) = \frac{1}{2} e^{((y-\frac{1}{2}) \cdot (\beta^T \cdot x) - w \cdot \frac{(\beta^T \cdot x)^2}{2})} \cdot P(w)$$

- In [Polson et al. 2013], it is shown that

$$\int_0^\infty P(w, y|\beta, x) dw = \frac{e^{y\beta^T \cdot x}}{1 + e^{\beta^T \cdot x}}$$



# Poly-Gamma Augmentation (Contd)

- Let  $W$  be the set of Poly-Gamma variables  $w$  for  $(x, y) \in \mathcal{D}$

$$P(\beta | \mathcal{D}) \propto \int P(W, \mathcal{D} | \beta) \cdot P(\beta) dW$$

- Approximate the augmented joint distribution  $P(W, \mathcal{D} | \beta) \cdot P(\beta)$

$$\prod_{(x,y) \in \mathcal{D}} \frac{1}{2} e^{((y-\frac{1}{2}) \cdot (\beta^T \cdot x) - w \cdot \frac{(\beta^T \cdot x)^2}{2})} \cdot P(w) \cdot \prod_i N(s_i | \mu_i, \sigma_s) \cdot \prod_j N(t_j | c_j, \sigma_t)$$

# Gibbs Sampling Algorithm

- Conditional distribution of  $w_i$

$$\begin{aligned} P(w_i | \beta, \mathbf{D}) &\propto e^{-w_i \frac{(\beta^T \cdot x_i)^2}{2}} P(w_i) \\ &= PG(w_i | 1, \beta^T \cdot x_i) \end{aligned}$$

- Conditional distribution for  $\beta_j$

$$\begin{aligned} P(\beta_j | \beta_{-j}, W, \mathbf{D}) &\propto \prod_{(x_i, y_i) \in \mathbf{D}} \prod_{x_{ij} \neq 0} \frac{1}{2} e^{\left( (y_i - \frac{1}{2}) \cdot (\beta^T \cdot x_i) - w_i \cdot \frac{(\beta^T \cdot x_i)^2}{2} \right)} N(\beta_j | \mu_{\beta_j}, \sigma_{\beta_j}) \\ &= N(\beta_j | m_j, V_j) \end{aligned}$$

$$m_j = V_j \left( \frac{\mu_{\beta_j}}{\sigma_{\beta_j}^2} + \sum_{(x_i, y_i) \in \mathbf{D}} \sum_{x_{ij} \neq 0} \left( (y_i - \frac{1}{2}) \cdot x_{ij} + w_i \cdot x_{ij} \cdot \sum_{l \neq j} \beta_l \cdot x_{il} \right) \right)$$

$$\frac{1}{V_j} = \frac{1}{\sigma_{\beta_j}^2} + \sum_{(x_i, y_i) \in \mathbf{D}} \sum_{x_{ij} \neq 0} w_i \cdot x_{ij}^2$$

# Predictive Distribution

- Let  $r$  samples be drawn from the posterior of  $\beta$
- Let the  $l^{\text{th}}$  sample be  $\beta^l = (s_1^l, \dots, s_c^l, t_1^l, \dots, t_p^l)$

$$P(y_{ij} = \text{fit} | \mathbf{D}) = \int P(y_{ij} = \text{fit} | \beta) \cdot P(\beta | \mathbf{D}) d\beta$$
$$\approx \frac{1}{r} \sum_{l=1}^r \frac{1}{1 + e^{\alpha(s_i^l - t_j^l) - b_1}} \cdot \frac{1}{1 + e^{-\alpha(s_i^l - t_j^l) + b_2}}$$

# Experimental Results

- Consider 6 real-life shoes datasets with between 10M and 33M transactions
- Baseline model
  - Product size  $t_j = c_j$
  - Customer size  $s_i =$  Average size of products purchased by customer
  - Logistic regression model with feature  $(s_i - t_j)$  to predict outcome
- Bayesian Logit model
  - Predict outcome with highest probability
- Performance metric: weighted AUC
- Results (% improvements over baseline)

Dataset	A	B	C	D	E	F
Bayesian	17.71	18.28	19.7	25.78	20.22	19.42

# Leveraging Customer and Product Features

- Means of latent size priors are obtained by performing regression over customer ( $f_i$ ) and product ( $g_j$ ) features [AC 2009]

- **Generative model:**

for each customer  $i$ ,

draw latent size  $s_i \sim \mathcal{N}(w_f \cdot f_i, \sigma_s^2)$

for each product  $j$ ,

draw latent size  $t_j \sim \mathcal{N}(w_g \cdot g_j, \sigma_t^2)$

for each transaction  $(i, j, y_{ij}) \in T$ ,

select  $y_{ij} = \textit{small}$  with probability  $P(y_{ij} = \textit{small} | .)$

select  $y_{ij} = \textit{fit}$  with probability  $P(y_{ij} = \textit{fit} | .)$

select  $y_{ij} = \textit{large}$  with probability  $P(y_{ij} = \textit{large} | .)$

- Perform least squares regression to learn parameters  $w_f$  and  $w_g$  from customer and product size samples

# Incorporating Customer Personas

- Latent size for persona  $k$  of customer  $i$ :  $s_{ik}$
- Latent variable containing persona involved in transaction  $(i, j, y_{ij})$ :  $z_{ij}$

- **Generative model:**

for each customer  $i$ ,

draw persona distribution  $\theta_i \sim \text{Dir}(\alpha)$

for each persona  $k$  draw latent size  $s_{ik} \sim N(w_f \cdot f_k, \sigma_s^2)$

for each product  $j$ ,

draw latent size  $t_j \sim N(w_g \cdot g_j, \sigma_t^2)$

for each transaction  $(i, j, y_{ij}) \in T$ ,

draw persona  $z_{ij} \sim \text{Mult}(\theta_i)$

select  $y_{ij} = \text{small}, \dots$  with probability  $P(y_{ij} = \text{small} | z_{ij}), \dots$

- Gibbs Sampling algorithm can be extended to draw  $z_{ij}$  samples

# Summary

- Learning semantically rich representations critical for future AI applications, several ML techniques
  - Conversational systems, content summarization, video metadata generation
  - Deep Learning, Probabilistic Models, Tensor Factorization
- **Deep Learning** to learn embeddings
  - Allows semantic matching between questions and snippets
  - Loss functions like Siamese loss that aim to maximize difference in class scores perform better
- **Probabilistic Graphical Models** to learn latent sizes
  - Priors handle data sparsity, prevent overfitting
  - Posteriors model uncertainty in data
  - Leverage all the available signals