

Machine Learning and AI at Etsy

March 7, 2018

Liangjie Hong
Head of Data Science, Etsy Inc.

Liangjie Hong

- **Head of Data Science at Etsy** since Aug, 2016
- **Senior Manager of Research at Yahoo Research** in Sunnyvale, CA
Leading science efforts for personalization and search sciences.
- Published papers in SIGIR, WWW, KDD, CIKM, AAI, WSDM, RecSys and ICML (2400+ citations)
- WWW 2011 Best Poster Paper Award
WSDM 2013 Best Paper Nominated
RecSys 2014 Best Paper Award
- Program committee members in KDD, WWW, SIGIR, WSDM, AAI, EMNLP, ICWSM, ACL, CIKM, IJCAI and various journal reviewers
- PhD in Machine Learning from Lehigh University

Etsy



OUR MISSION

Reimagine commerce in
ways that build a more
fulfilling and lasting world

Etsy – A Global Marketplace



Artifact Bags
Omaha, NE

Photo by: Dana Damewood and Jackie Sterba



Clap Clap
Los Angeles, CA

Photo by: Bert Youn and Mimi Kim



redravenstudios
Pittsburgh, PA

Photo by: Janelle Bendycki



Little Hero Capes
Somerset, MA

Photo by: Rich Vintage Photography



Cattails Woodwork
Hermitage, PE, Canada

Photo by: Cattails Woodwork



Room for Emptiness
Berlin, Germany

Photo by: Room for Emptiness



sukrachand
Brooklyn, NY

Photo by: sukrachand



Nicole Porter Design
Saint Paul, MN

Photo by: Nicole Porter Design



noemiah
Montreal, QC, Canada

Photo by: noemiah



Lorgie
Fremantle, WA, Australia

Photo by: Lorgie



Jeremiah Collection
San Francisco, CA

Photo by: Matthew Reamer



Docksmith
Brunswick, ME

Photo by: Docksmith



purlBKnit
Brooklyn, NY

Photo by: purlBKnit



Julia Astreou
Nicosia, Cyprus

Photo by: Panagiotis Mina



Moira K. Lime
Omaha, NE

Photo by: Moira K. Lime



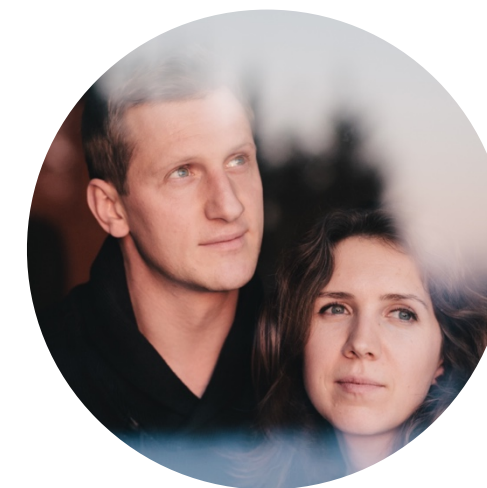
Nested Yellow
Portland, OR

Photo by: Jessica Dremov and Nested Yellow



Habitables
Madrid, Spain

Photo by: Habitables



Woodstorming
Kaunas, Lithuania

Photo by: Ilona & Martynas from Instudija



karoArt
Dublin, Ireland

Photo by: Christine Burns



ADIKILAV
Jerusalem, Israel

Photo by: Shlomit Koslowe



My A La Mode Boutique
Ecuador

Photo by: My A La Mode Boutique

Etsy – A Global Marketplace

What can you sell on Etsy?



Handmade Goods



Vintage

(20 years or older)



Craft Supplies

By The Numbers

1.9M

active sellers

31.7M

active buyers

\$2.8B

annual GMS

45+M

items for sale



AI Application: E-commerce

The image displays three overlapping e-commerce search results pages:

- Amazon (Top Left):** Search results for "liszt". It features a sidebar with categories like "CDs & Vinyl", "Books", and "Digital Music". A featured product is "LISZT Consolations and Liebesträume For the Piano" by Franz Liszt, available as a paperback for \$7.52 or a Kindle edition for \$9.99.
- eBay (Top Right):** Search results for "camera". It shows 2,241,207 results. A blue tooltip says "Quickly compare offers in our new grouped view". The page includes filters for "Categories" (Cameras & Photo, Digital Cameras, etc.) and "Shop by Category".
- Etsy (Bottom):** Search results for "wabi sabi". It shows 6,213 results across various categories like "Jewelry & Accessories", "Clothing & Shoes", and "Home & Living". Specific items include a "Kintsugi bowl" for \$84.41, a "BIGFOOT Bowl" for \$42.00, a "Wabi-sabi Oversize Clutch bag" for \$65.00, and a "Wabi-Sabi definition, dictionary art print" for \$4.99.

AI Application: E-commerce

- **Search**
- **Recommendation**
- **Advertising**

AI Application: E-commerce

- **Search**
 - **Recommendation**
 - **Advertising**
-

- **Shopping**
- **Discovery**

AI Application: E-commerce



AI Application: E-commerce

- **Search**
 - Generic search v.s. E-commerce search
 - Relevance, Revenue, Diversity, Discovery

AI Application: E-commerce

- **Search**
 - Generic search v.s. E-commerce search
 - Relevance, Revenue, Diversity, Discovery
- **Recommendation**
 - Rating/favorite prediction
 - Clicks and purchase funnel
 - Revenue, Seasonal, Occasion, Inventory

AI Application: E-commerce

- **Search**

- Generic search v.s. E-commerce search
- Relevance, Revenue, Diversity, Discovery

- **Recommendation**

- Rating/favorite prediction
- Clicks and purchase funnel
- Revenue, Seasonal, Occasion, Inventory

- **Advertising**

- Two-sided marketplace

AI Application: E-commerce

- **Search**
 - **Recommendation**
 - **Advertising**
-

- **How to measure**
- **How to optimize**

AI Application: E-commerce

The image shows two screenshots of the Etsy website. The top screenshot is a search for "wabi sabi" with 6,213 results. The bottom screenshot is a search for "jewelry box" with 241,017 results. Both screenshots show the Etsy navigation bar, category filters, search filters, and product listings.

Top Screenshot: Search for "wabi sabi"

- Search bar: wabi sabi
- Search button: Search
- Navigation: Sell on Etsy, Home, Favorites, You, Cart
- Categories: Jewelry & Accessories, Clothing & Shoes, Home & Living, Wedding & Party, Toys & Entertainment, Art & Collectibles, Craft Supplies & Tools, Vintage
- Filters: wabi sabi art, wabi sabi ceramics, wabi sabi bowl, wabi sabi pottery, wabi sabi necklace, wabi sabi jewelry
- Sort by: Relevancy
- Left sidebar filters:
 - Special offers: On sale
 - All categories: Home & Living, Art & Collectibles, Jewelry, Craft Supplies & Tools
 - Shipping: Free shipping, Ready to ship in 1 business day, Ready to ship within 3 business days
 - Shop location: Anywhere, United States, Custom
- Product listing: Kintsugi bowl, kintsugi ceramic er, KanelaSuri, \$84.41, Only 1 available, 5 stars (77)

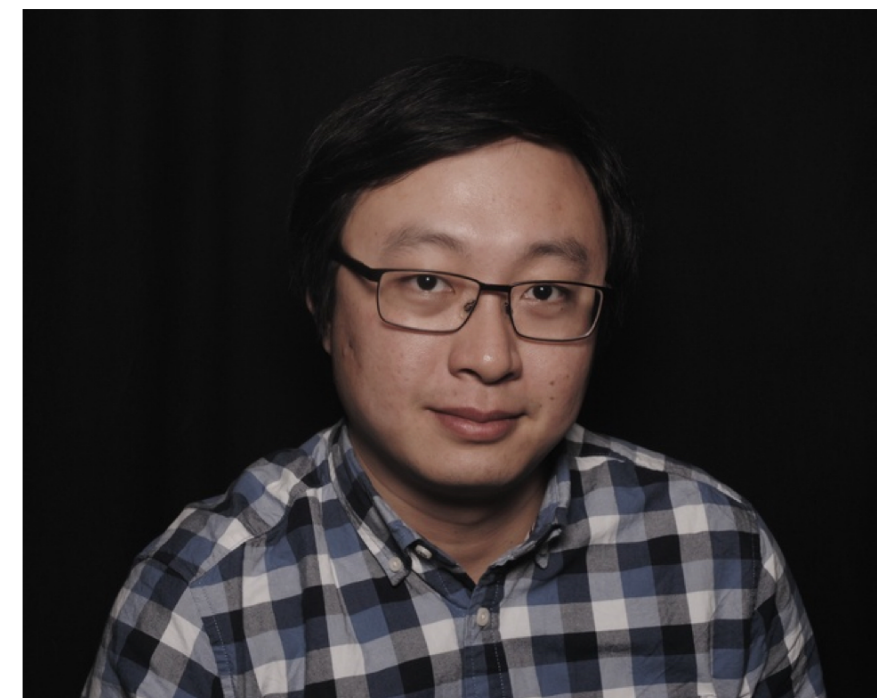
Bottom Screenshot: Search for "jewelry box"

- Search bar: jewelry box
- Search button: Search
- Navigation: Sell on Etsy, Home, Favorites, You, Cart
- Categories: Jewelry & Accessories, Clothing & Shoes, Home & Living, Wedding & Party, Toys & Entertainment, Art & Collectibles, Craft Supplies & Tools, Vintage
- Filters: jewelry box wood, wooden jewelry box, large jewelry box, small jewelry box, jewelry box vintage, personalized jewelry box
- Sort by: Relevancy
- Left sidebar filters:
 - All categories: Jewelry, Home & Living, Craft Supplies & Tools, Weddings
 - Shipping: Free shipping, Ready to ship in 1 business day, Ready to ship within 3 business days
 - Special offers: On sale
 - Shop location: Anywhere, United States, Custom
- Product listings:
 - Raven box, handmade boxes, steamp... ST3jewellery, 5 stars (35), \$30.95
 - Bridesmaid Gift / Popular Bridesmaid... SugarAndChicShop, 5 stars (1,208), \$45.00
 - Matte Black Custom Branded Laserc... lzrbeams, 5 stars (162), \$85.00
 - Personalized Memory Box, Keepsake ... EngraveMyMemories, 5 stars (6,548), \$29.95, Eligible orders get 10% off

AI Application: E-commerce

Optimizing Gross-Merchandise-Value (GMV) in E-commerce Search

- **Liang Wu**, PhD Student from Arizona State University
- **Diane Hu**, Staff Data Scientist at Etsy
- **Liangjie Hong**, Head of Data Science at Etsy



AI Application: E-commerce

Optimizing Gross-Merchandise-Value (GMV) in E-commerce Search

$$GMV = \sum_{\underbrace{\forall s \in S}_{\text{A search session}}} \sum_{\underbrace{\forall i^s}_{\text{An item in s}}} \underbrace{Price(i^s)}_{\text{Price of } i^s} \underbrace{Pr(\Phi = 1 | i^s, q^s)}_{\text{Prob of purchase}},$$

AI Application: E-commerce

Optimizing Gross-Merchandise-Value (GMV) in E-commerce Search

- **Purchase Decision Process**

The screenshot shows a search page for "rosy wedding dress" with 72 results. Four products are highlighted with callouts to their respective product pages:

- Product 1:** "60 Colors Chiffon Rosy Long Party Dress Evenin...". Price: \$51.50. Callout: "Rosy brown dress chiffon party dress rosy brown prom dress chiffon cocktail dress bow back dress rosy brown bridesmaid dresses chiffon dress". Price: \$39.00.
- Product 2:** "Ivory Mauve Flower Girl Dress - Flower girl Dress...". Price: \$69.00. Callout: "Ivory Mauve Flower Girl Dress - Flower girl Dress Rosy Mauve - Flower Girl Dress - Dress for Flower Girls - flower girls Pink Mauve". Price: \$69.00+.
- Product 3:** "Rosy Mauve Satin Bridal Sash - Rosy Mauve We...". Price: \$14.00.
- Product 4:** "Rosy Mauve Satin Bridal Sash - Rosy Mauve We...". Price: \$14.00.

Search Page

Product Page

AI Application: E-commerce

Optimizing Gross-Merchandise-Value (GMV) in E-commerce Search

- **Click Decision(s) from Search-Result-Page (SERP)**
- **Purchase Decision(s) from Listing Page**


$$Pr(\Phi = 1|i, q) = \underbrace{Pr(\Psi = 1|i, q)}_{\text{click model}} \underbrace{Pr(\Phi = 1|\Psi = 1, i, q)}_{\text{purchase model}},$$

AI Application: E-commerce

Optimizing Gross-Merchandise-Value (GMV) in E-commerce Search

- Click Decision(s) from Search-Result-Page (SERP)

$$NDCG_K(q) = N_{max}^{-1} \sum_{r=0}^{K-1} \frac{2^{l(r^{-1})}}{\log(1+r)},$$


$$\mathcal{L}_c = N_{max}^{-1} \sum_{i=1}^m \frac{2^{l(i)}}{\log(1 + \sum_{i_b=1, i_b \neq i_a}^m \sigma(f_c(x_a) - f_c(x_b)))},$$

f_c is learned by a neural-network model through back-prop.

AI Application: E-commerce

Optimizing Gross-Merchandise-Value (GMV) in E-commerce Search

- **Purchase Decision from Listing Page**

$$\mathcal{L}_p = \sum_{i=1}^N Price(i) \log\{1 + \exp[-l'_i(w_p x_i)]\} + \|w_p\|^2,$$

Price-Weighted Logistic Regression

AI Application: E-commerce

Optimizing Gross-Merchandise-Value (GMV) in E-commerce Search

Sessions	Queries	Items	Avg. Items per Session
334,931	239,928	6,347,251	19.0
Keywords	Buyers	Sellers	Avg. Items per Query
631,778	270,239	550,025	26.5

AI Application: E-commerce

Optimizing Gross-Merchandise-Value (GMV) in E-commerce Search

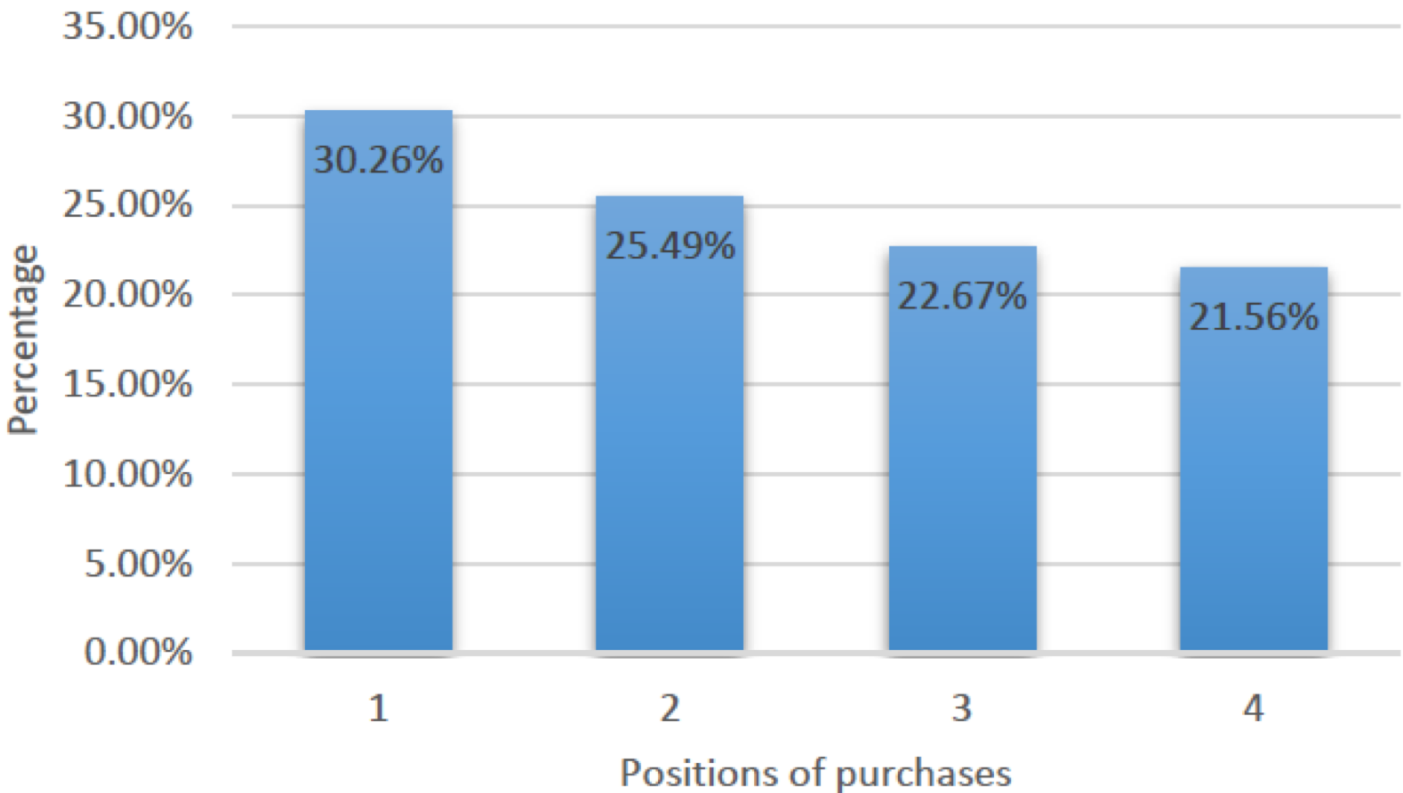


Figure 2: Position distribution of items being purchased in the top 4 spots of a search result page. The first position achieves the most purchases, while nearly 70% of purchases are in the lower positions.

AI Application: E-commerce

Optimizing Gross-Merchandise-Value (GMV) in E-commerce Search

Relevance	Low Level	Sum of TF
		Sum of $\log TF$
		Sum of Normalized TF
		Sum of \log Normalized TF
		Sum of IDF
		Sum of $\log IDF$
		Sum of ICF
		Sum of $TF-IDF$
		Sum of $\log TF-IDF$
		$TF-\log IDF$
		$Length$
		$\log Length$
	High Level	BM_{25}
		$\log BM_{25}$
LM_{DIR}		
LM_{JM}		
LM_{ABS}		
Revenue	$Price$	
	$Price - Cat.Mean$	
	$(Price - Cat.Mean)/Cat.Mean$	

Click	RankNet [1]	RNet
	RankBoost [10]	RBoost
	AdaRank [39]	ARank
	LambdaRank [2]	LRank
	ListNet [3]	LNet
	MART [12]	MART
Purchase	LambdaMART [38]	LMART
	SVM [4]	SVM
	Logistic Regression [28]	LR
Both	Random Forest [22]	RM
	Weighted Purchase [44]	WT
	LMART+RM	LMRM
	LETORIF	LETORIF

AI Application: E-commerce

Category	Method	Click NDCG@5			Purchase NDCG@5			Revenue NDCG@5		
		Train	Vali	Test	Train	Vali	Test	Train	Vali	Test
Click	RNet	0.1743	0.1731	0.1378**	0.1672	0.1721	0.1676**	0.1692	0.1700	0.1356**
	RBoost	0.2150	0.1768	0.1323**	0.2150	0.1768	0.1715**	0.2150	0.1768	0.1311**
	ARank	0.1718	0.1711	0.1351**	0.1718	0.1711	0.1706**	0.1718	0.1711	0.1358**
	LRank	0.1694	0.1688	0.1360**	0.1678	0.1711	0.1672**	0.1713	0.1719	0.1366**
	LNet	0.1665	0.1703	0.1355**	0.1601	0.1682	0.1620**	0.1646	0.1696	0.1348**
	MART	0.2700	0.1758	0.1380**	0.2155	0.1803	0.1796*	0.2696	0.1688	0.1408**
	LMART	0.3056	0.1777	0.1412	0.3056	0.1777	0.1717**	0.3056	0.1777	0.1370**
Purchase	SVM	0.1785	0.1772	0.1336**	0.1831	0.1754	0.1755**	0.1816	0.1752	0.1320**
	LR	0.1978	0.1739	0.1310**	0.1978	0.1739	0.1782**	0.1978	0.1739	0.1332**
	RM	0.3359	0.1698	0.1363**	0.3329	0.2305	0.1798**	0.3327	0.1685	0.1376**
Both	WT	0.1970	0.1682	0.1334**	0.1815	0.1763	0.1761**	0.1781	0.1648	0.1375**
	LMRM	0.2943	0.2597	0.1354**	0.3087	0.2530	0.1688**	0.2943	0.2594	0.1332**
	LETORIF	0.1765	0.1550	0.1351**	0.2731	0.1841	0.1801	0.2039	0.1698	0.1494

Symbol * indicates that the method is outperformed by the best one by 0.05 statistical significance level, ** indicates 0.01.

AI Application: E-commerce

Category	Method	Rev@1	Rev@2	Rev@3	Rev@4	Rev@5	Rev@6	Rev@7	Rev@8	Rev@9	Rev@10
Click	RNet	4.47**	4.69**	4.89**	4.91*	5.06**	5.23**	5.21**	5.33**	5.46**	5.55**
	RBoost	4.57**	4.69**	4.69**	4.76**	4.97**	5.17**	5.23**	5.36**	5.49**	5.57**
	ARank	4.37**	4.66**	4.76**	4.90**	5.06**	5.20*	5.33**	5.47**	5.59**	5.67**
	LRank	4.38**	4.61**	4.74**	4.86**	5.07**	5.25**	5.42**	5.42**	5.67**	5.78**
	LNet	4.30**	4.59**	4.78**	4.99**	5.16**	5.35**	5.49**	5.61**	5.63**	5.63**
	MART	4.62	4.72**	4.86**	5.04**	5.26**	5.47**	5.47**	5.64**	5.74**	5.86**
	LMART	4.46*	4.54**	4.73**	5.10**	5.31**	5.56**	5.75**	5.90*	6.01**	6.14**
Purchase	SVM	4.41**	4.54**	4.76**	4.77**	4.95**	5.16**	5.34**	5.50**	5.64**	5.77**
	LR	4.29**	4.65**	4.65**	4.69**	4.74**	4.81*	4.94**	4.97**	5.11**	5.11**
	RM	4.52**	4.82**	4.86**	5.02**	5.18**	5.33*	5.50**	5.66**	5.79**	5.92**
Both	WT	4.52**	4.69**	4.80**	4.85**	5.01**	5.07**	5.23**	5.32**	5.35**	5.41**
	LMRM	4.42**	4.50**	4.72**	5.08**	5.23**	5.41**	5.57**	5.60**	5.73**	5.85**
	LETORIF	4.58**	4.90	5.08	5.47	5.64	5.85	6.02	6.19	6.40	6.54

Symbol * indicates that the method is outperformed by the best one by 0.05 statistical significance level, ** indicates 0.01.

Application: E-commerce

Optimizing Gross-Merchandise-Value (GMV) in E-commerce Search

- This work is about optimizing GMV in Session
 - How about long-term GMV?
 - How about other discovery?
 - ...
- First step in optimizing user engagements in E-commerce search.

AI in E-commerce

AI in E-commerce at Etsy

- Multi-modal Deep-learning based Search Solution (KDD 2016)
- Probabilistic Graphical Model based Personalization Recommendation (KDD 2014)
- Ensemble Learning based CTR Prediction Solution (AdKDD 2017/KDD 2017)
- Buzzsaw: A System for High Speed Feature Engineering (SysML 2018)



