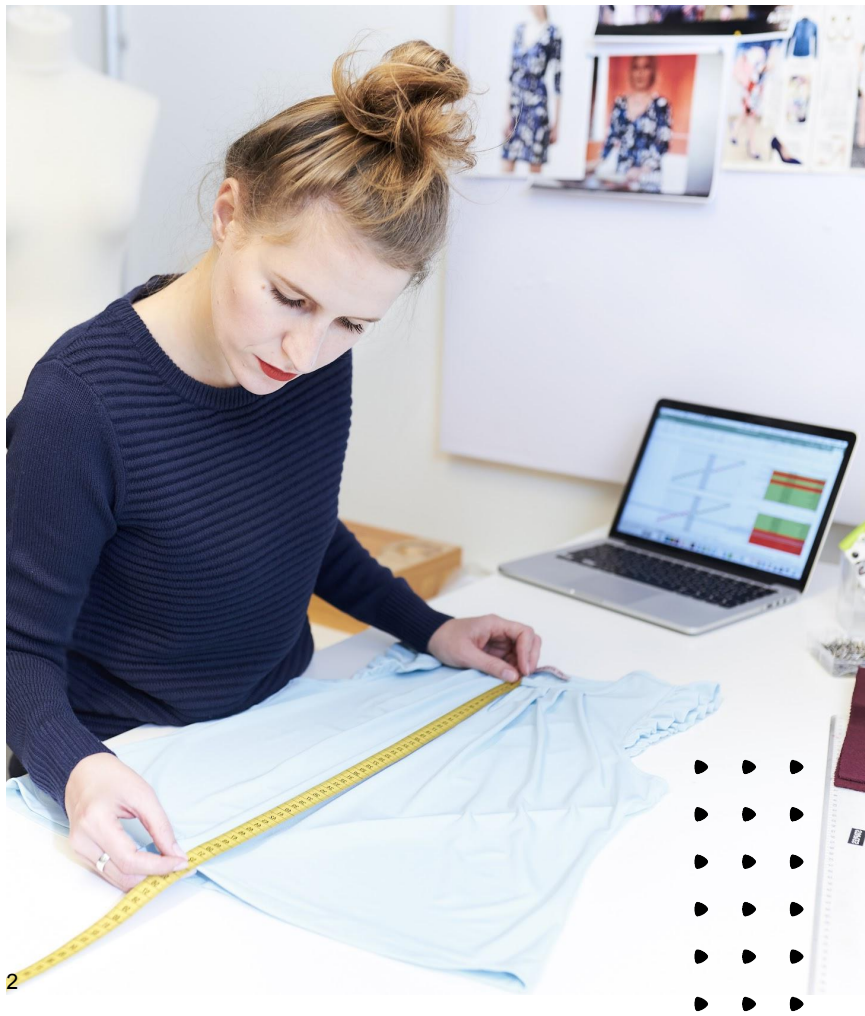


Knowing When You Don't Know in Online Fashion: An Uncertainty Aware Size Recommendation Framework

Hareesh Bahuleyan, Julia Lasserre, Leonidas Lefakis and Reza Shirvany





Outline


- Personalized Size Recommendation
- Dataset Noise Challenges
- Key Contributions
 - Measuring Prediction Uncertainty
 - Uncertainty-Aware Training
- Experiments and Results
- Summary and Conclusions

Personalized Size Recommendation Task

Help and contact Free delivery and returns 100-day return policy


Women **Men** Kids

Get the Look Clothing Shoes Sports Accessories Care Designer Brands **Sale %** Pre-owned Search




Wrangler
GOOD TIMES TEE - Print T-shirt
33,95 € VAT included
★★★★☆ 2


Colour: blue graphite



We recommend size L. [Why?](#)


L (Only 1 left) ▾

Add to bag 

 1-2 working days
PLUS premium delivery **Free**

★★★★☆ 2


Colour: blue graphite



We recommend size L. [Why?](#)

L (Only 1 left) ▾

Add to bag

 1-2 working days
PLUS premium delivery **Free**

Personalized Size Recommendation Task

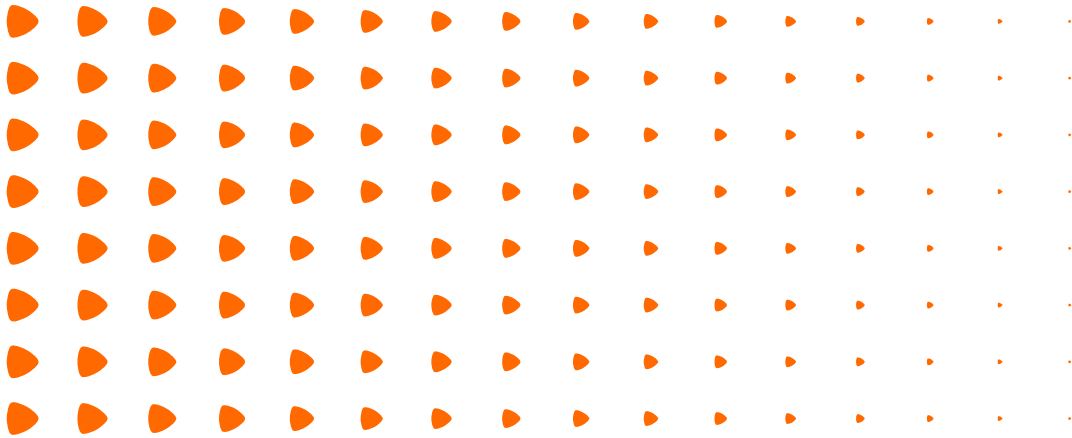
- Supervised learning problem:

Each training instance consists of:

- Support purchases: sequence of articles previously purchased by the customer
 - Query article
 - Ground truth size of the query article
-
- Test time: Predict the size for a new query article
-
- Deep Learning Model: [Attention Gets You the Right Size and Fit in Fashion](#) [RecSys 2020: FashionXRecsys Workshop]

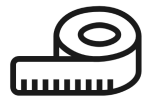
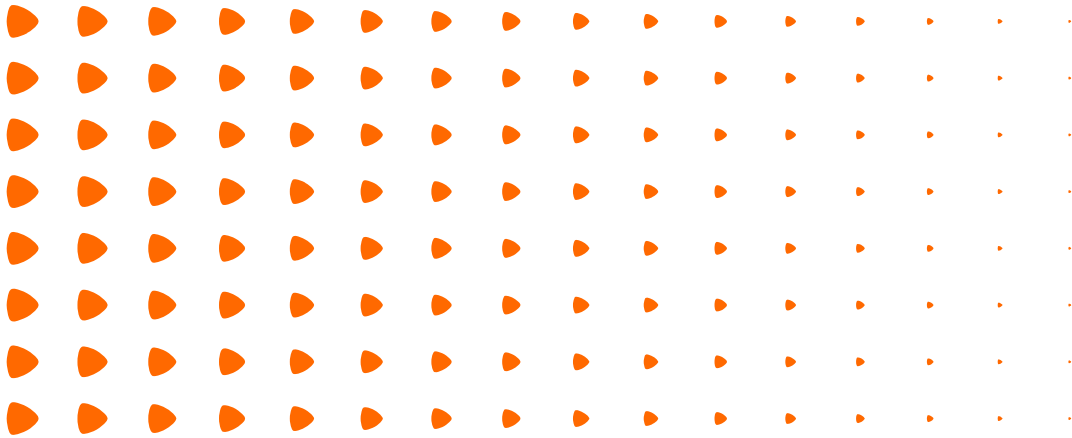
Dataset Noise Challenges

- Style and fit preferences can differ among customers
- Differences in sizing systems across countries, brands, etc.
- Vanity sizing to boost customer esteem
- Multiple users behind the same customer account
- And more ...



Key Contributions

- ❑ Uncertainty Quantification for Size and Fit Prediction
- ❑ Uncertainty-Aware Training



Measuring Prediction Uncertainty

Uncertainty Based Decision Making

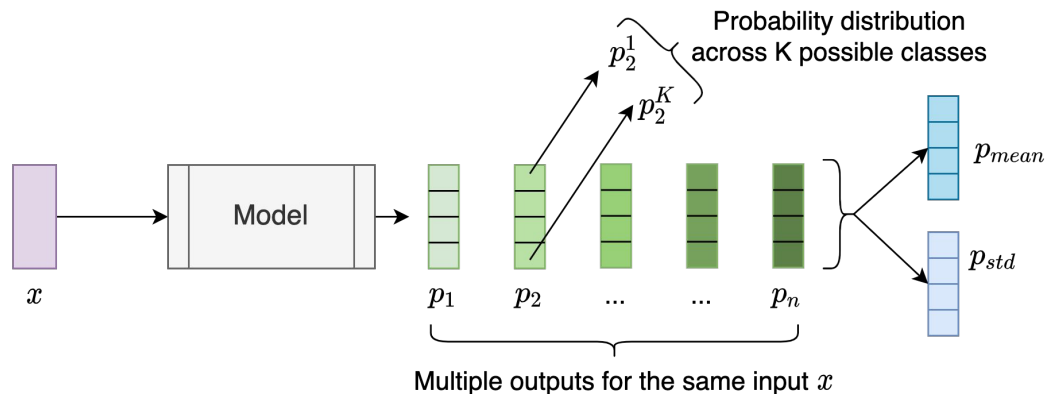
- When interacting with customers it is important to know the degree of uncertainty associated with our ML model predictions
- Might be better to take an alternate action in case of high uncertainty
- Enables building trust with customers → positive feedback loop where they interact more with our systems



Uncertainty Quantification Metrics

- Maximum Softmax Probability (`max_proba`)
 - Probability assigned to the predicted size
- Predictive Entropy (`entropy`)
 - Shannon entropy of the output softmax distribution
- And more ...

Uncertainty Quantification Metrics (Monte-Carlo Dropout)



- Monte-Carlo Dropout at Test Time [[Gal et al. \(2016\)](#)]
- Normalized Standard Deviation of the *Predicted Size* (`stddev_pred`)
- Normalized Standard Deviation of the *True Size* (`stddev_true`)

Comparison of Uncertainty Metrics

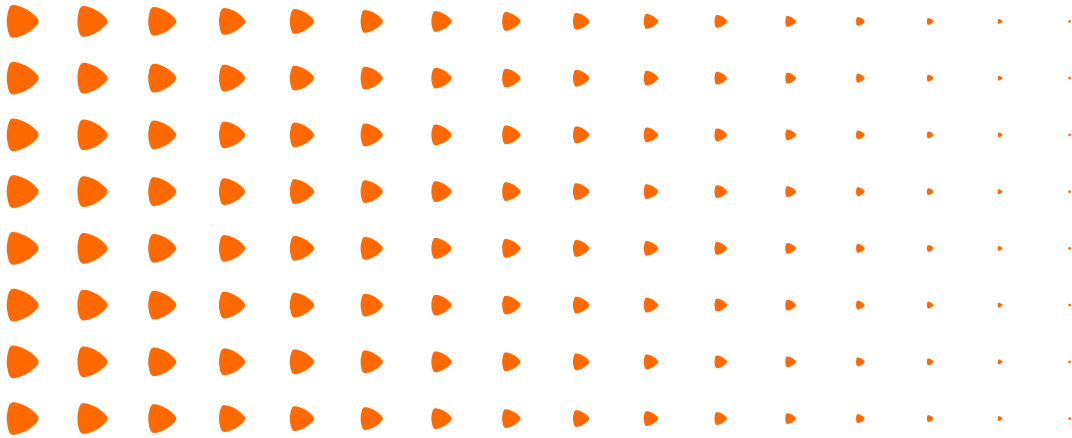
- Size Prediction Accuracy at Different Test Set Coverage Levels

	Coverage Level	
Uncertainty Metric	80%	60%
max_proba	66.92	72.28
entropy	66.72	71.57
stddev_pred	65.47	70.34

Comparison of Uncertainty Metrics

- Size Prediction Accuracy at Different Coverage Levels

	Coverage Level	
Uncertainty Metric	80%	60%
max_proba	66.92	72.28
entropy	66.72	71.57
stddev_pred	65.47	70.34
stddev_true	74.74	90.01



Uncertainty Aware Training Two Approaches

Approach 1: Hard Weighting by Dataset Pruning

- Four stages:
 1. K-fold training
 2. Uncertainty estimation based on `stddev_true`
 3. Dataset pruning (of most uncertain instances)
 4. Re-training on the cleaner subset
- Decide on percentage of dataset pruned

Approach 2: Soft Weighting by Instance Re-weighting

- Three stages:
 1. K-fold training
 2. Uncertainty estimation based on `stddev_true`

Approach 2: Soft Weighting by Instance Re-weighting

- Three stages:
 1. K-fold training
 2. Uncertainty estimation based on `stddev_true`
 3. Retraining with instance reweighting

$$\text{ReweightedLoss} = \exp(-\beta * \sigma_i) * \text{CrossEntropyLoss}$$

where σ_i refers to `stddev_true` and β is a hyperparameter

Experiment and Results

- In-house dataset of 5.7M purchases by 260k customers spanning across 60+ garment categories (upper/lower garments, shoes), 2000+ brands and 1000+ sizes

	Coverage Level	80%	60%
Baseline (Full Noisy Dataset)		66.92	72.28

Experiment and Results

- In-house dataset of 5.7M purchases by 260k customers spanning across 60+ garment categories (upper/lower garments, shoes), 2000+ brands and 1000+ sizes

	Coverage Level	80%	60%
Baseline (Full Noisy Dataset)		66.92	72.28
Hard-Weighting	max_proba	66.90	72.44
	stddev_pred	67.26	72.66
	stddev_true	67.11	72.30

Experiment and Results

- In-house dataset of 5.7M purchases by 260k customers spanning across 60+ garment categories (upper/lower garments, shoes), 2000+ brands and 1000+ sizes

	Coverage Level	80%	60%
Baseline (Full Noisy Dataset)		66.92	72.28
Hard-Weighting	max_proba	66.90	72.44
	stddev_pred	67.26	72.66
	stddev_true	67.11	72.30
Soft-Weighting	max_proba	66.83	72.21
	stddev_pred	67.27	72.52
	stddev_true	67.59	73.04

Experiment and Results

- Comparison to other competitive baselines

	Coverage Level	80%	60%
Baseline (Full Noisy Dataset)		66.92	72.28
Soft-Weighting	Deep Abstaining Recommender	67.02	72.60
	Confident Learning	67.76	73.04
	stddev_true	67.59	73.04

- Confident Learning - quadratic complexity in the number of classes

Summary and Conclusions

- Noisy data challenges in personalized size recommendation
- Explored and compared uncertainty metrics
 - `stddev_true` is a strong indicator of prediction uncertainty
- Uncertainty-aware retraining
 - dataset pruning
 - instance level reweighting
- Future work:
 - circumvent multi-stage training
 - apply proposed methodology on other domains

Knowing When You Don't Know in Online Fashion: An Uncertainty Aware Size Recommendation Framework



Join Zalando Machine Learning Community


<https://pages.beamery.com/zalando/page/machinelearning>

Join us!


About us.

This is Zalando's Machine Learning community. You may know us as a fashion e-commerce platform, but not many know about our 40+ machine learning projects, research, and products that are running behind the scenes.

By signing up, you'll get the latest on our current machine learning products and updates on some of our [open source projects](#) and datasets.




Why join?




Tech News

We'll send you digests of our research publications, open source projects, product launches, and tech blog articles.



Conferences

Find us at the next conference, or get exclusive invites to our own organized networking events.




Vacancies

We'll send you all vacancies in our machine learning teams that are tailored to your level and skillsets.

Behind the scenes.

Take a look at some of our Personalization teams and how they work tirelessly to apply machine learning to solve complex fashion problems for our customers. We're a very serious bunch.



Applied Scientist - Fashion Recommendations

Job ID: 2301504 | mis à jour le 9/23/2020

Applied Scientist - Machine Learning & DL

Job ID: 2352248 | mis à jour le 9/24/2020

Senior Principal Scientist - Artificial Intelligence (w/m/d)

Job ID: 2196437 | mis à jour le 9/23/2020

Senior Applied Scientist - Logistics Algorithms (W/M/D)

Job ID: 2336235 | mis à jour le 9/23/2020

Principal Applied Scientist - Payments (W/M/D)

Job ID: 2313398 | mis à jour le 9/23/2020

Thanks!

jobs.zalando.com

22

October, 2021

Third Workshop on Recommender Systems in Fashion