




zalando

Outfit Generation and Recommendation – an Experimental Study

MARJAN CELIKIK
MATTHIAS KRIMSE
TIMO DENK
PIERRE GAGLIARDI
SAHAR MBARECK
DUY PHAM

Sep 26, 2020

INTRODUCTION

- **Outfit**: set of pairwise compatible fashion items

- **Outfit generation**:

- **Personalized outfit generation**:

CONTENTS

- MOTIVATION
- NON-PERSONALIZED ALGORITHMS
- ALGORITHMS FOR PERSONALIZED OUTFIT GENERATION
- EMPIRICAL EVALUATION

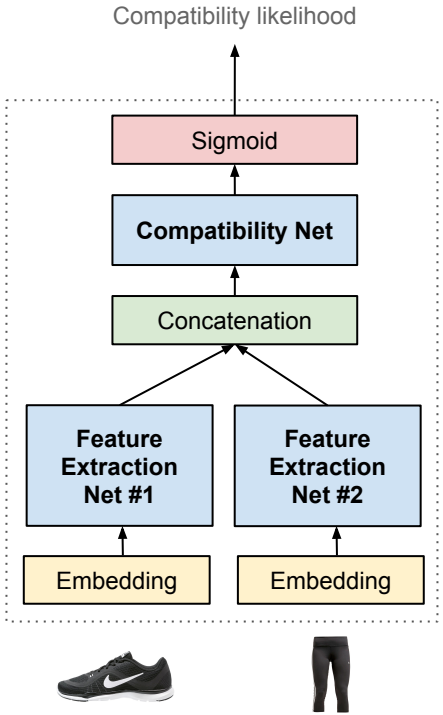
MOTIVATION

- **No extensive performance analysis** of different outfit generation algorithms in the fashion domain
- We contribute to close this gap by **comparison of 3 main classes** of algorithms
  - Siamese-Nets based
  - LSTM-based
  - Attention-based
- In addition, we **adapt** a few of these algorithms to the personalized outfit generation use-case

ALGORITHMS FOR ITEM COMPATIBILITY

NON-PERSONALIZED OUTFIT GENERATION

SIAMESE NETS FOR OUTFIT GENERATION

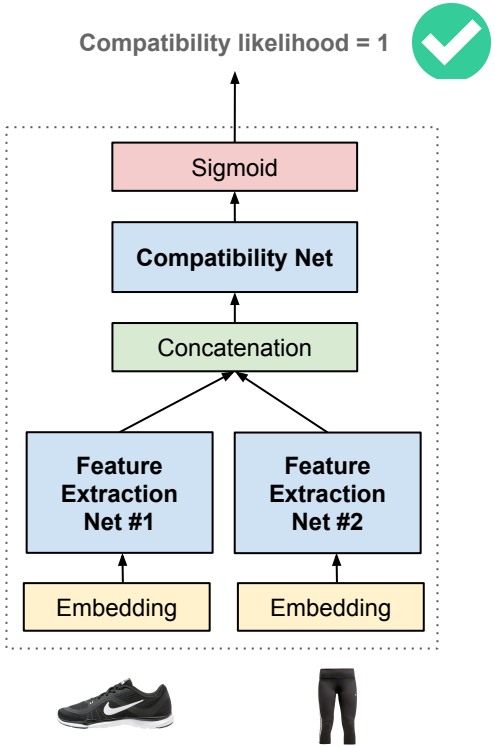


Learns a **compatibility function** which outputs a score “how compatible a pair of items is”

Consists of identical **feature extraction blocks** and **compatibility block**

Contrastive training on compatible pairs of items versus pairs with random replacements

SIAMESE NETS FOR OUTFIT GENERATION

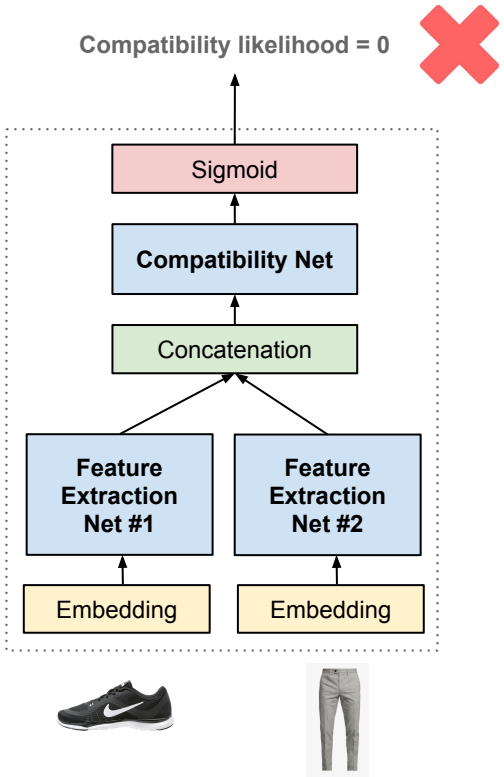


Learns a **compatibility function** which outputs a score for “how compatible a pair of items is”

Consists of identical **feature extraction blocks** (without weight sharing) and **compatibility block**

Contrastive training on compatible pairs from stylist created outfits and pairs with random replacements

SIAMESE NETS FOR OUTFIT GENERATION

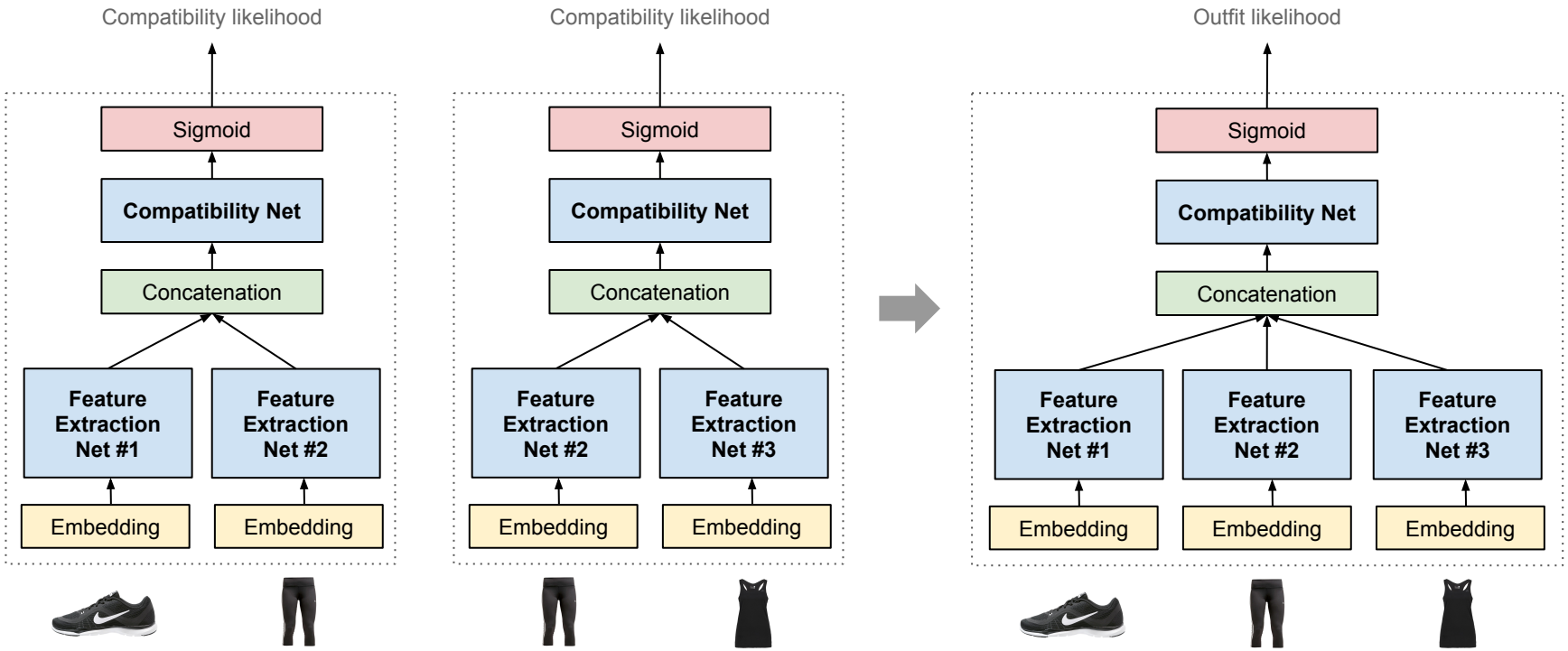


Learns a **compatibility function** which outputs a score for “how compatible a pair of items is”

Consists of identical **feature extraction blocks** (without weight sharing) and **compatibility block**

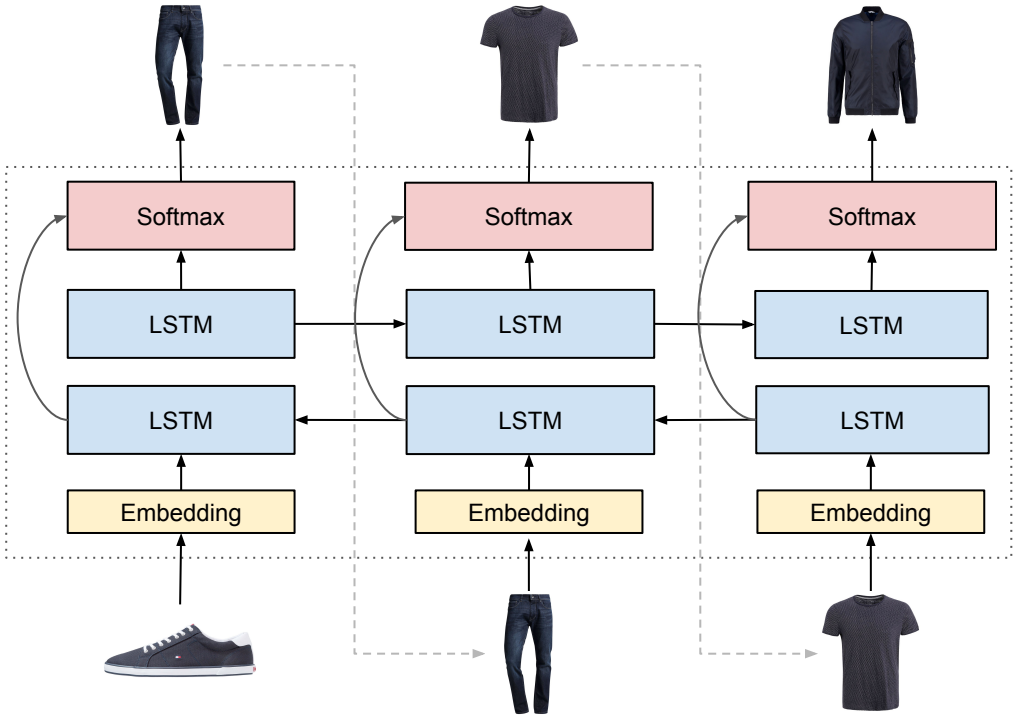
Contrastive training on compatible pairs from stylist created outfits and pairs with random replacements

SIAMESE NETS FOR OUTFIT GENERATION



zalando

LSTMs FOR OUTFIT GENERATION

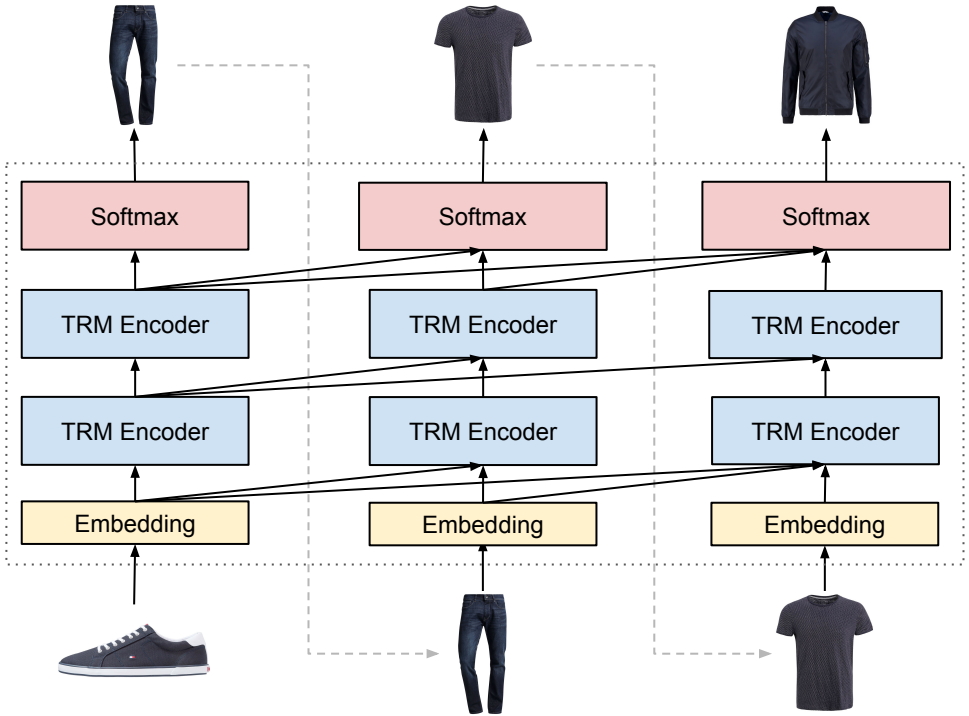


Models item compatibility via learning the **transition probability** between items as a proxy

Outfit represented as a **sequence** of fashion categories

Trained by **predicting the next** and **the previous item** in the outfit

GPT FOR OUTFIT GENERATION

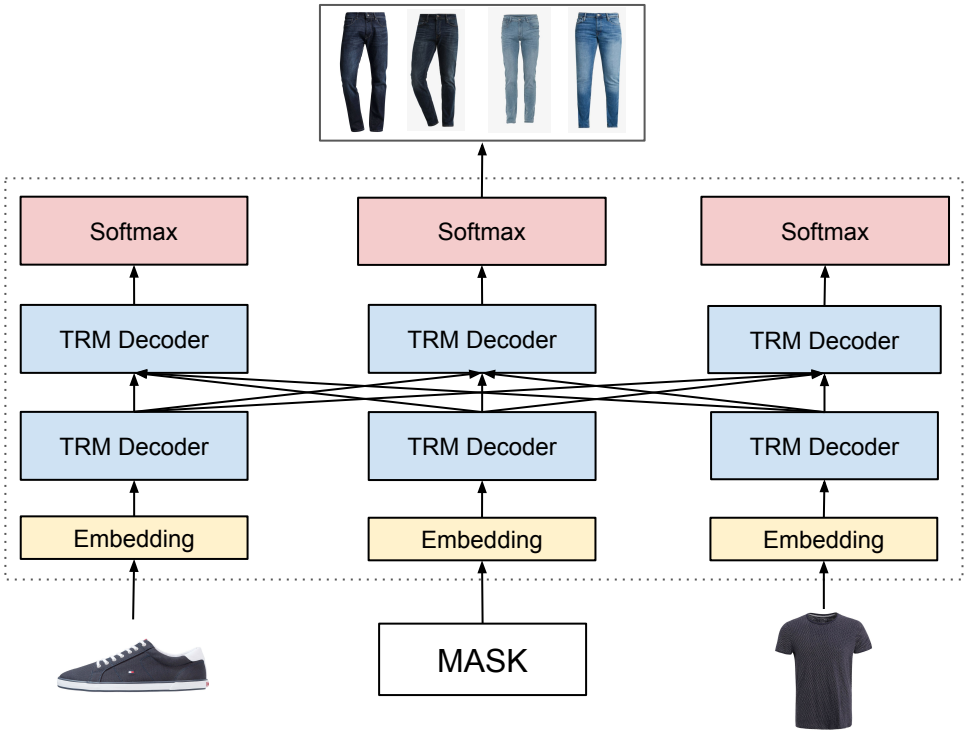


Item compatibility modeled via **self-attention**, where each item attends to all previous items

Outfit **represented as a set** by removing positional encoding

Model trained by **predicting the next item in the outfit**

BERT FOR OUTFIT GENERATION



Item compatibility modeled via **self-attention** - each item attends to all other items

Outfit r**epresented as a set** by removing the positional encoding as well as the next sentence task

Model trained by **filling in the blank** of a masked item in the outfit

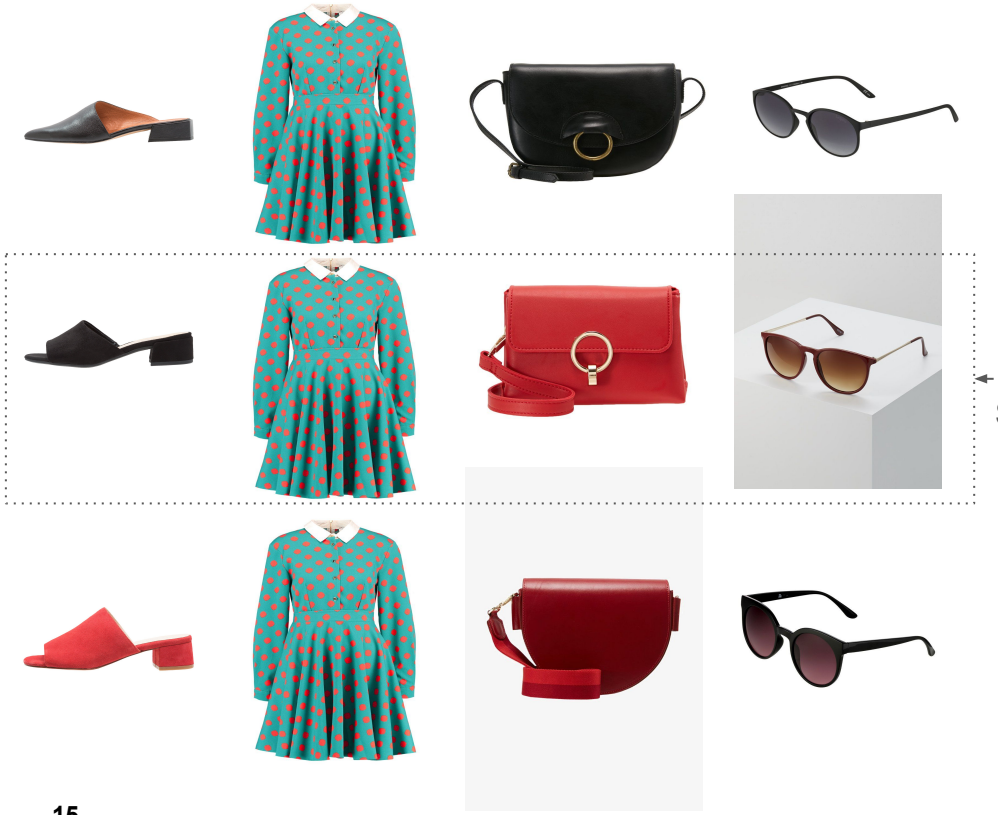
PERSONALIZED OUTFIT GENERATION

GENERIC ALGORITHM FOR PERSONALIZED OUTFIT GENERATION

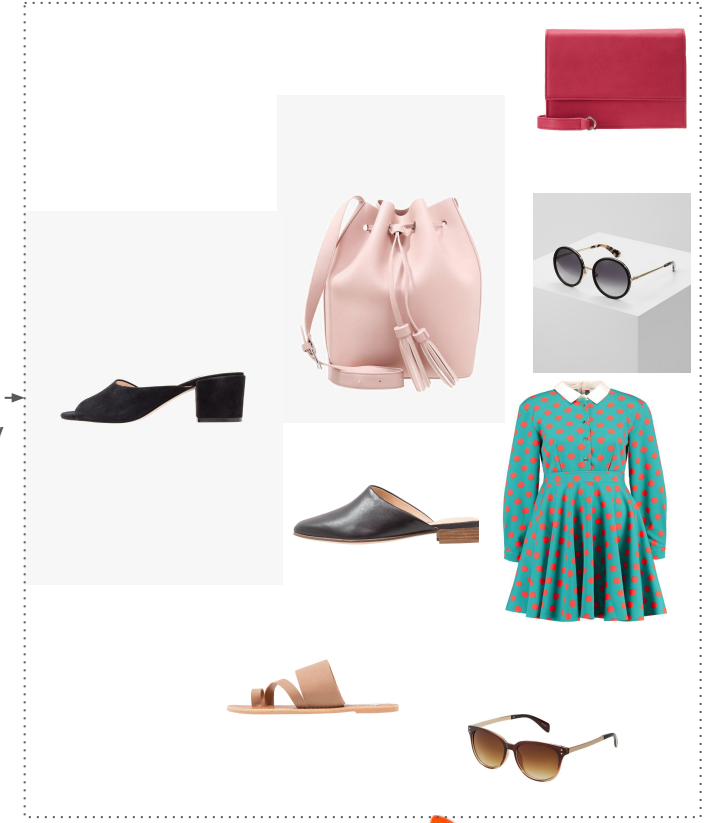
Generate enough outfits

zalando

GENERIC ALGORITHM FOR PERSONALIZED OUTFIT GENERATION



Similarity



Generate enough outfits

User click history

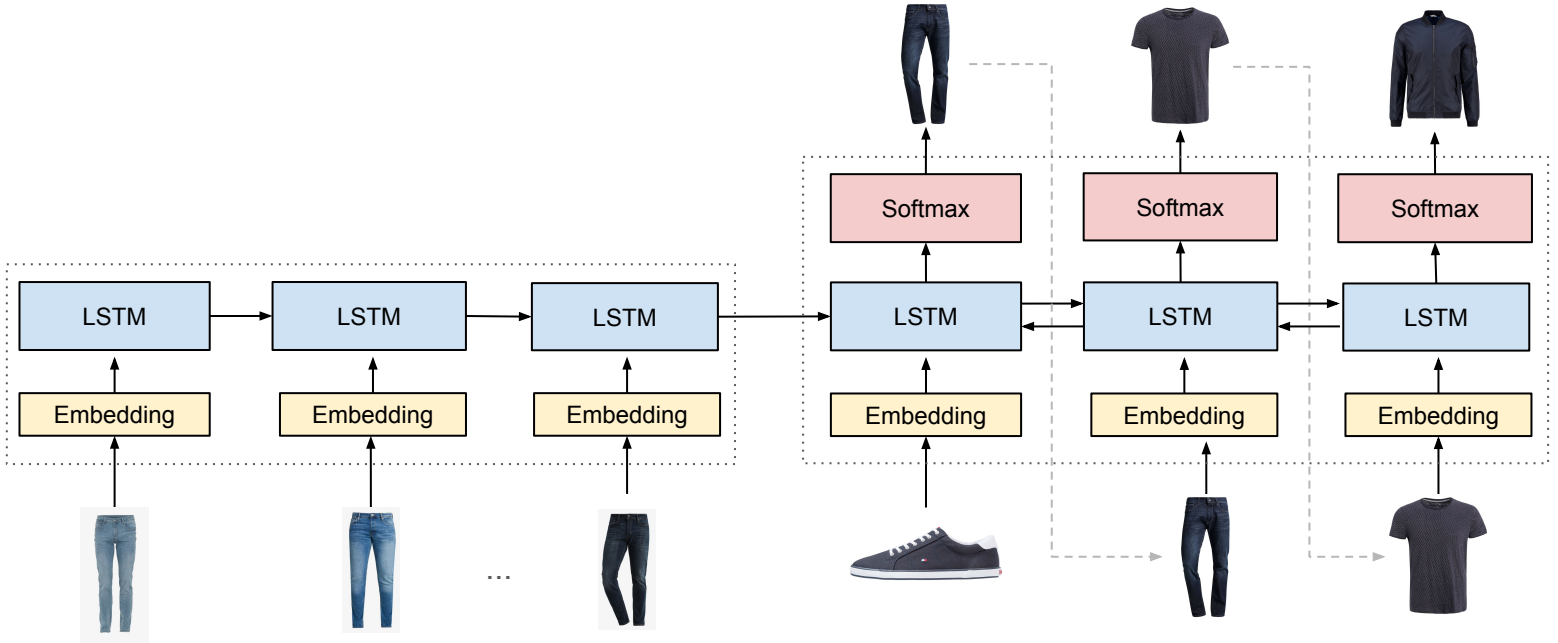
GENERIC ALGORITHM FOR PERSONALIZED OUTFIT GENERATION

Sort by similarity

Generate enough outfits

User click history

SEQ-TO-SEQ LSTMS FOR PERSONALIZED OUTFIT GENERATION

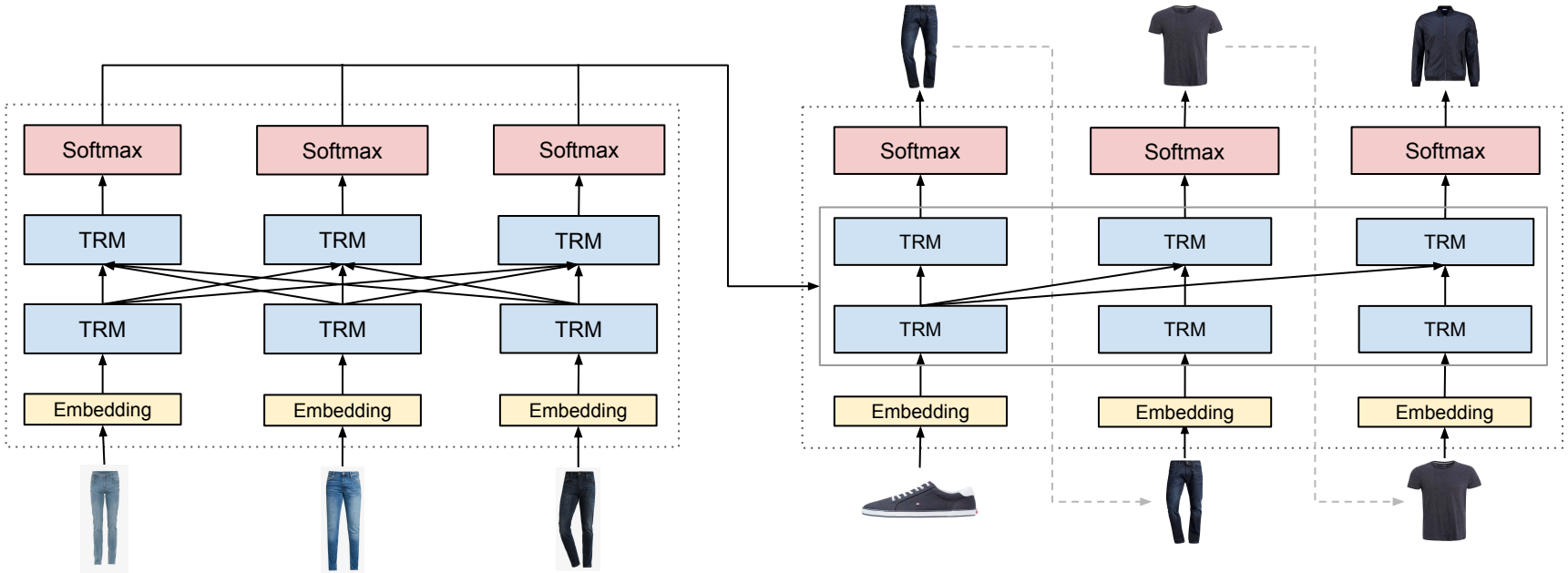


Learns a fixed dimensional representation of the **user interacted items**

Generates an outfit sequence conditioned on the state of the first LSTM

zalando

TRANSFORMER FOR PERSONALIZED OUTFIT GENERATION

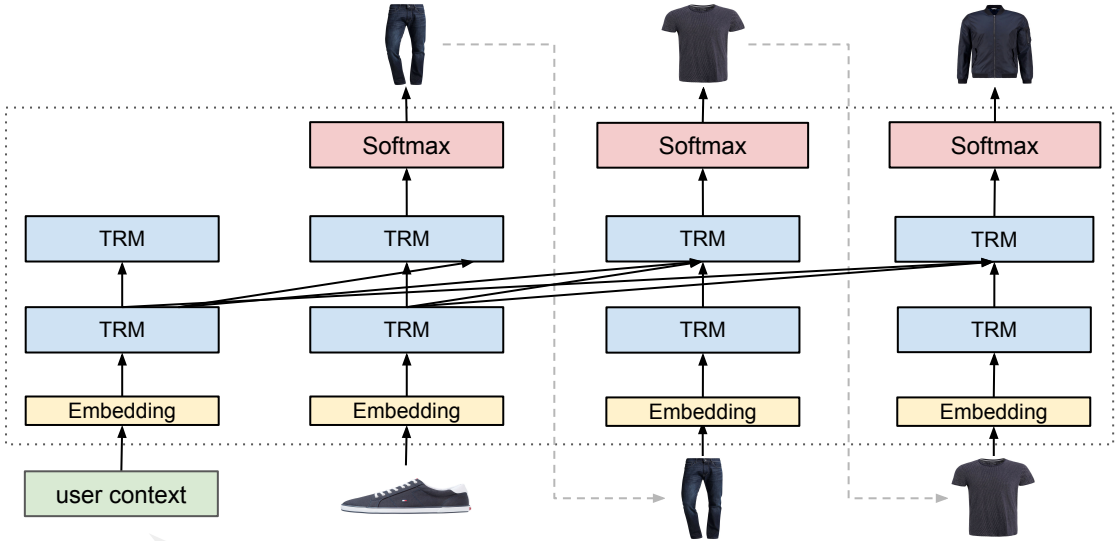


Encoder: learns a user preference signal from the interacted items, by attending each item to every other item

Decoder: generates an outfit by attending to the previous items and the user preference signal

zalando

CONTEXTUAL BERT AND GPT FOR PERSONALIZED OUTFIT GENERATION



Gender, age, location, favorite brands, favorite colors, height, body shape, etc.

Add **global contextual data** to BERT / GPT / TRM **as an additional token**

The fixed size context is **encoded as an embedding** with the same dimensionality as the item embedding

The model can attend to the context vector and utilize it for prediction as with any other token

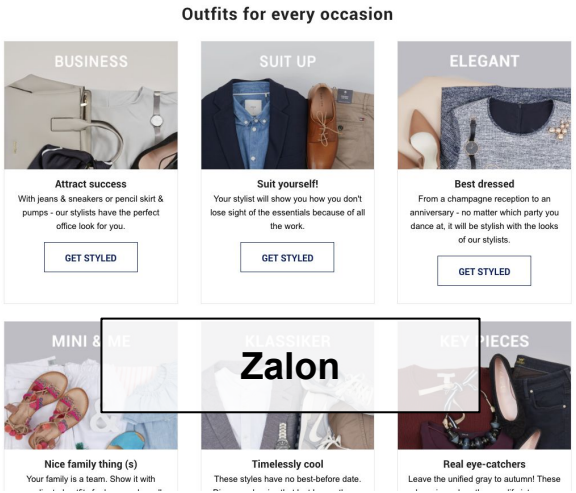
zalando

EMPIRICAL EVALUATION

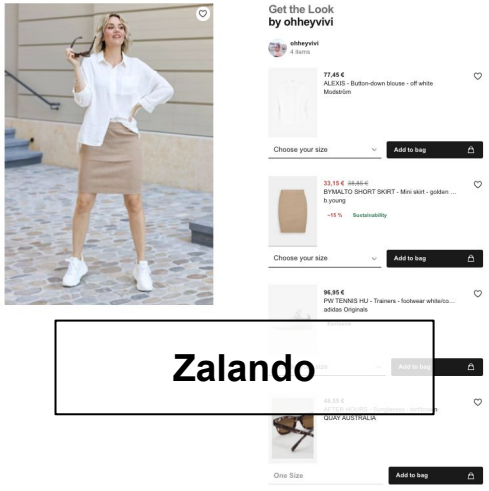
OUTFIT DATASETS

Name	Personalization	#Outfits	#Articles	Avg. outfit length
Zalon	Questionnaire	380k	31k[1]	5.0
Zalando GTL & STL	Click history	250k	65k[1]	4.5

[1] after filtering by occurrence frequency



Zalon



Zalando

zalando

RESULTS FOR NON-PERSONALIZED MODELS

Zalon Dataset			
Model	**Perplexity**	**Compatibility**	**FITB Accuracy**
Siamese	-	71.9%	0.1%
LSTM	28,637	64.1%	0.7%
GPT	**1,212**	**92.1%**	2.4%
BERT	9,934	89.0%	**4.8%**

Zalando Dataset			
Model	**Perplexity**	**Compatibility**	**FITB Accuracy**
Siamese	-	73.7%	0.4%
LSTM	34,290	68.6%	2.4%
GPT	**92**	96.9%	17.7%
BERT	182,586	**97.9%**	**49.3%**

Perplexity Metric

- $\exp(\mathrm{CE}(\mathrm{x}))$
- Captures the ability to autoregressively generate outfits
- GPT performs best

RESULTS FOR NON-PERSONALIZED MODELS

Zalon Dataset			
Model	**Perplexity**	**Compatibility**	**FITB Accuracy**
Siamese	-	71.9%	0.1%
LSTM	28,637	64.1%	0.7%
GPT	**1,212**	**92.1%**	2.4%
BERT	9,934	89.0%	**4.8%**

Zalando Dataset			
Model	**Perplexity**	**Compatibility**	**FITB Accuracy**
Siamese	-	73.7%	0.4%
LSTM	34,290	68.6%	2.4%
GPT	**92**	96.9%	17.7%
BERT	182,586	**97.9%**	**49.3%**

Compatibility Metric

- Accuracy at distinguishing invalid and real outfit
- Invalid outfits derived from real outfits by replacing one item randomly
- GPT and BERT are the best-performing

RESULTS FOR NON-PERSONALIZED MODELS

Zalon Dataset			
Model	**Perplexity**	**Compatibility**	**FITB Accuracy**
Siamese	-	71.9%	0.1%
LSTM	28,637	64.1%	0.7%
GPT	**1,212**	**92.1%**	2.4%
BERT	9,934	89.0%	**4.8%**

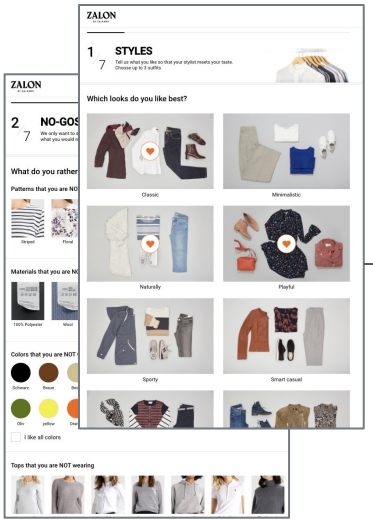
Zalando Dataset			
Model	**Perplexity**	**Compatibility**	**FITB Accuracy**
Siamese	-	73.7%	0.4%
LSTM	34,290	68.6%	2.4%
GPT	**92**	96.9%	17.7%
BERT	182,586	**97.9%**	**49.3%**

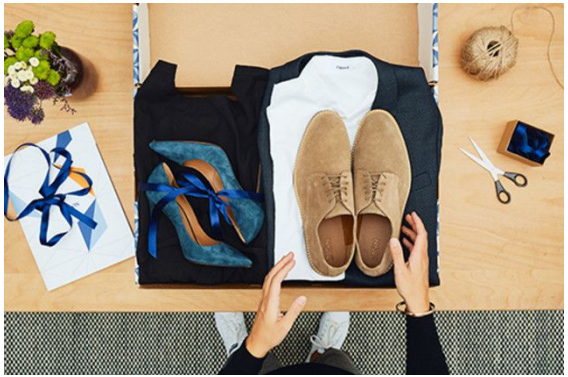
FITB accuracy

- Accuracy at predicting a masked-out item
- BERT performs best (FITB is its training task)

PERSONALIZED OUTFIT DATASETS: ZALON

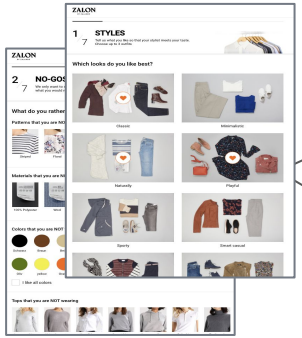
Name	Personalization	#Outfits	#Articles	Avg. outfit length
Zalon	Questionnaire	380k	31k	5.0
Zalando GTL & STL	Click history	250k	65k	4.5

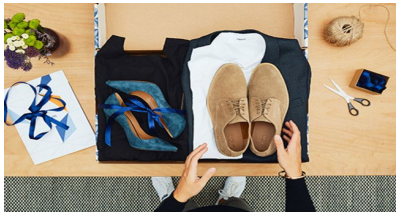




zalando

RESULTS FOR PERSONALIZED OUTFIT MODELS (ZALON)





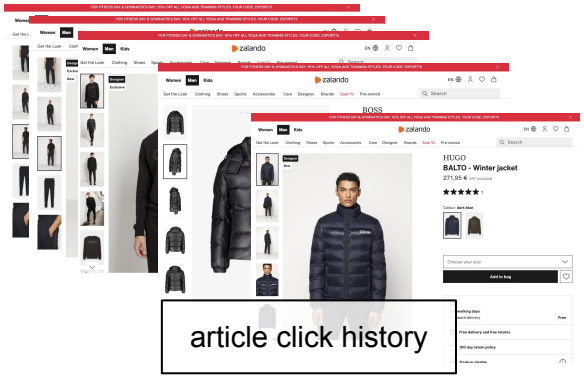
Stylist outfit
(physically shipped)

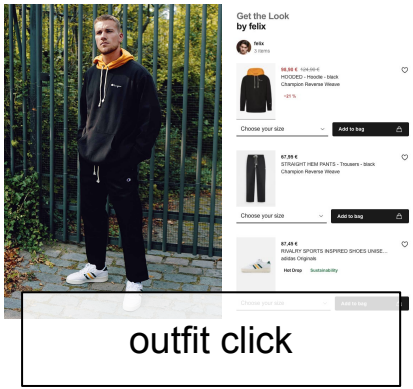
Outfit predictions

Personalized Zalon Dataset					
Model	**Brand-category KR**	**Color-category KR**	**Brand-color-category KR**	**Personalization rate**	**Item diversity rate**
Contextual GPT	**2.0%**	**2.3%**	**0.7%**	**0.5%**	5.6%
Contextual BERT	0.6%	1.6%	0.2%	**0.5%**	**33.6%**

PERSONALIZED OUTFIT DATASETS: ZALANDO

Name	Personalization	#Outfits	#Articles	Avg. outfit length
Zalon	Questionnaire	380k	31k	5.0
Zalando GTL & STL	Click history	250k	65k	4.5

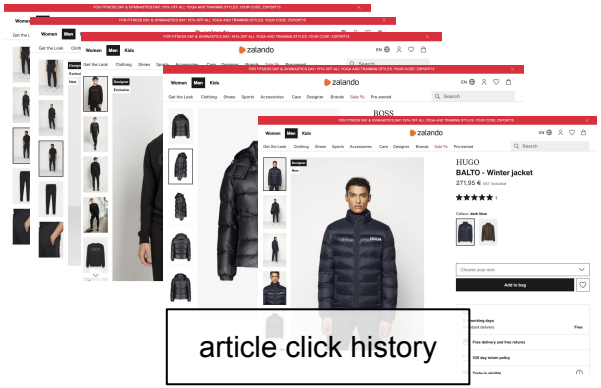


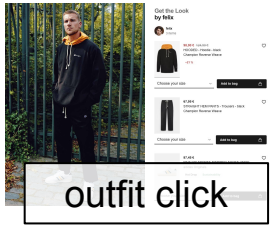


zalando

RESULTS FOR PERSONALIZED OUTFIT MODELS (ZALANDO)

Personalized Zalando Dataset					
Model	**Brand-category CTR**	**Color-category CTR**	**Brand-color-category CTR**	**Personalization rate**	**Item diversity rate**
Siamese Nets	5.8%	9.3%	2.7%	10.7%	7.7%
Transformer	**40.8%**	**40.2%**	**35.6%**	24.1%	31.4%
Seq-to-Seq LSTM	9.4%	12.8%	7.4%	**51.9%**	**35.7%**





zalando

SUMMARY

- **Comparison** of different algorithms for personalized and non-personalized outfit generation/recommendation
- Evaluation on **real-world data** from Zalando and Zalon
- **Transformer-based models** perform best