A Multimodal and Interpretable Recommender System for Massive Datasets

Barbara Rychalska

barbara.rychalska.dokt@pw.edu.pl

Supervisor: prof. dr hab. inż. Przemysław Biecek

24.05.2023

PhD Summary

- Industrial PhD done at Warsaw University of Technology and host company Synerise
- Supported by my NCN PRELUDIUM grant "MAD-NLP: Multiaspectual Diagnostics of NLP Systems"
- Research commercialized in 2 platforms.
- A part of the research was published as open source:
  - https://github.com/MI2DataLab/WildNLP
  - https://github.com/Synerise/cleora
- Research awarded with top places in international research competitions:

OGB-LSC @ KDD Cup 2021
Large - Scale Challenge

ACM RecSys Challenge

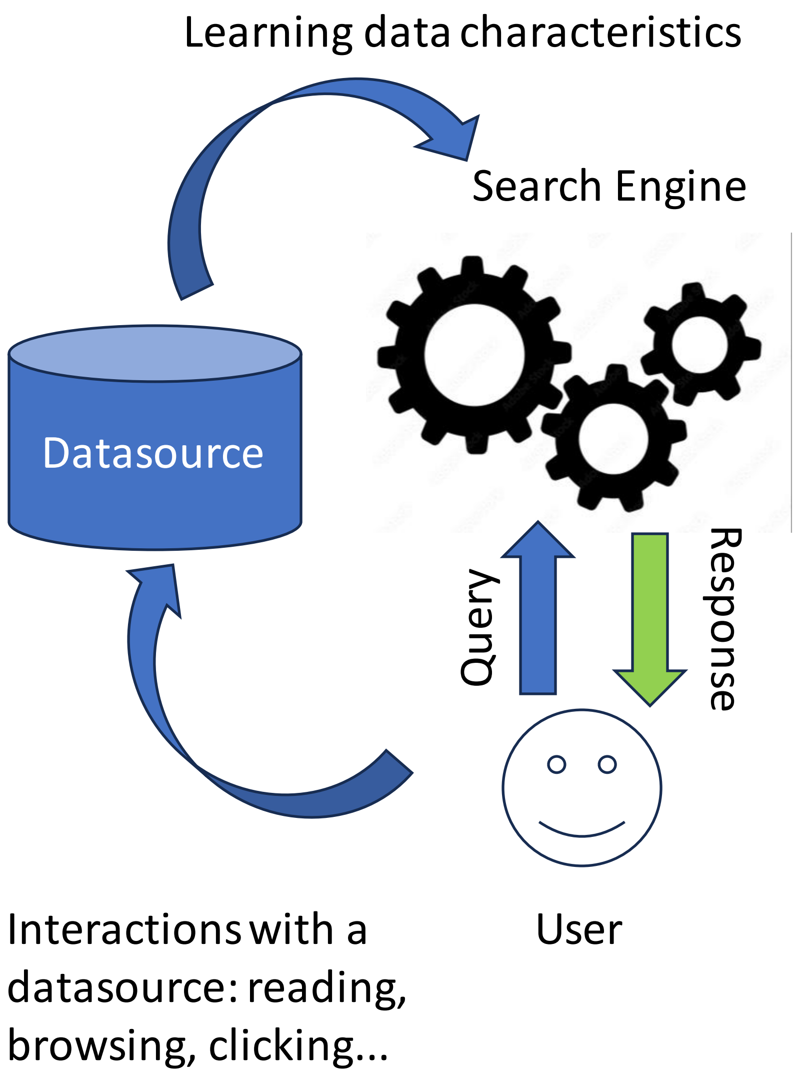
Rakuten

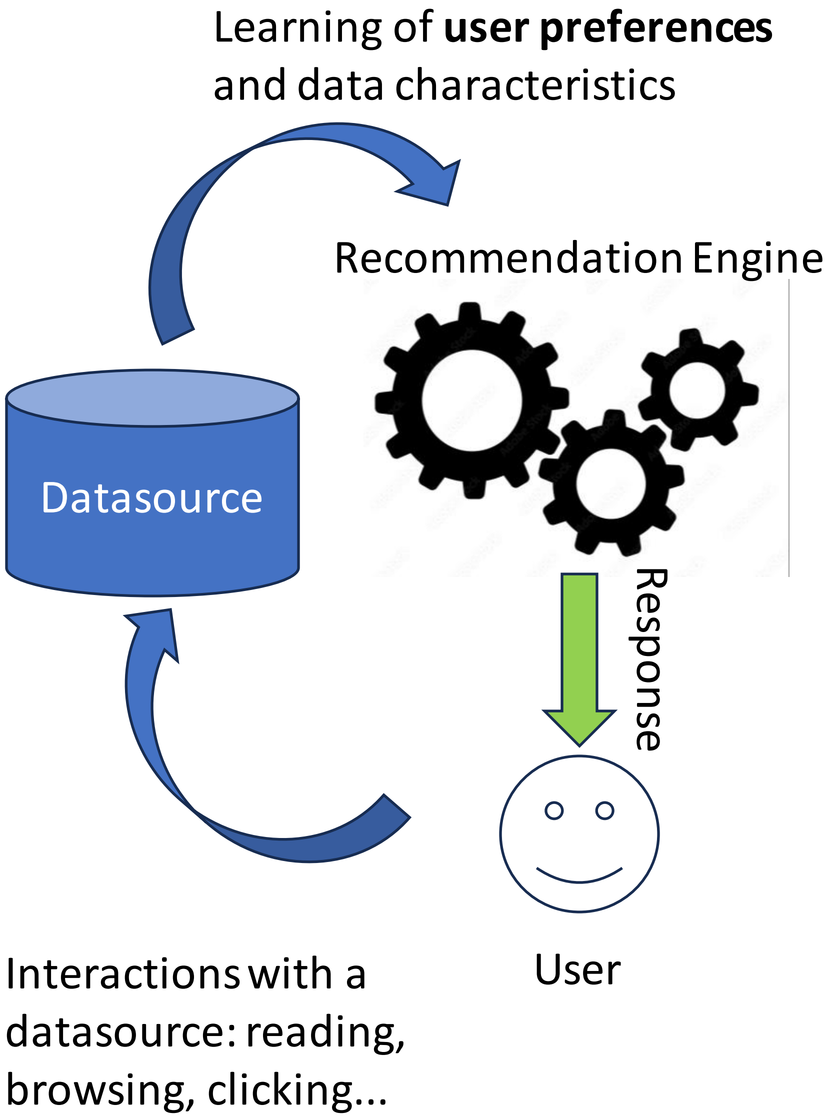
KDD Cup

Twitter RecSys Challenge

SIGIR eCom Rakuten Challenge

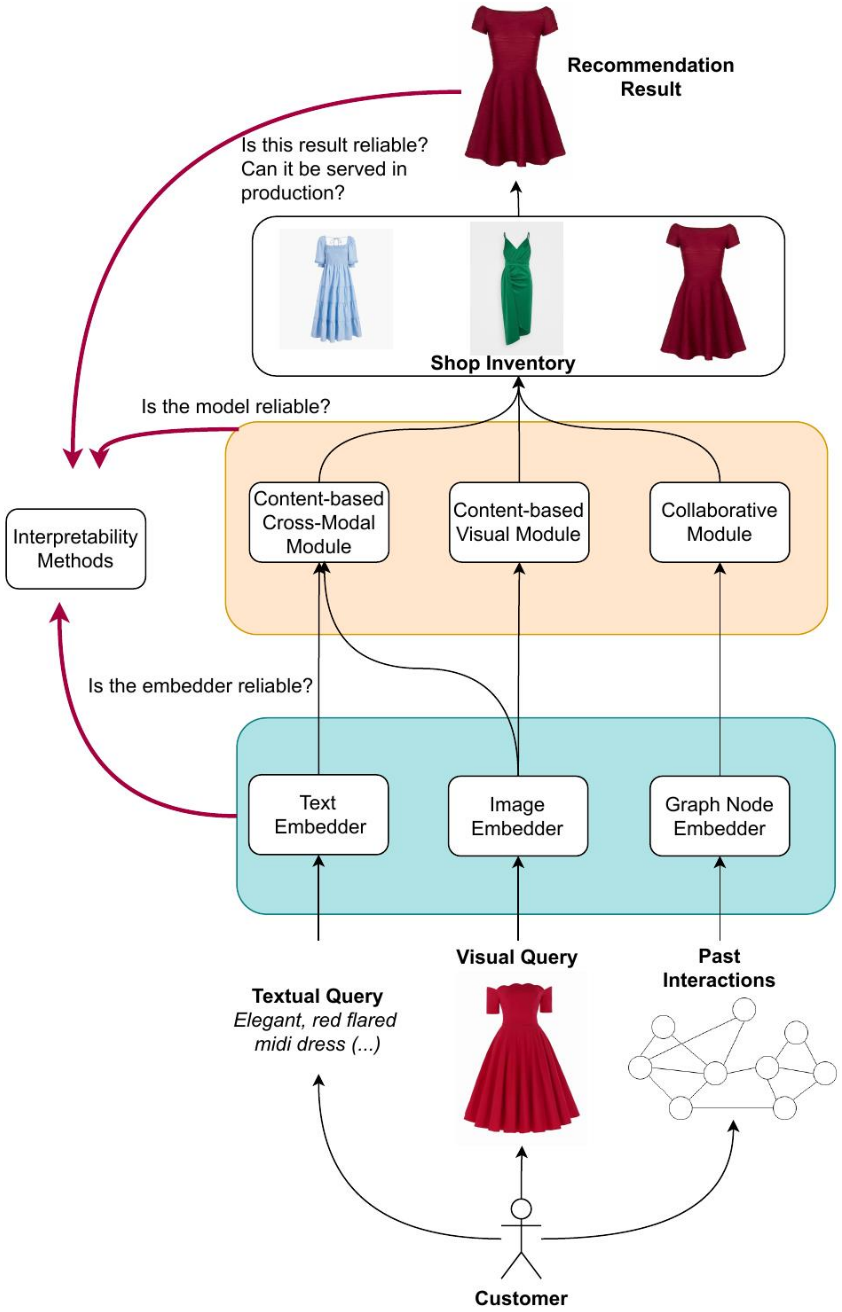
What is Recommendation?





A Multimodal and Interpretable Recommender System for Massive Datasets

- Multimodal Embedding Systems
- Collaborative Recommendation
- Content-based Uni-modal Recommendation
- Content-based Multimodal Recommendation
- Intepretability & Robustness Methods

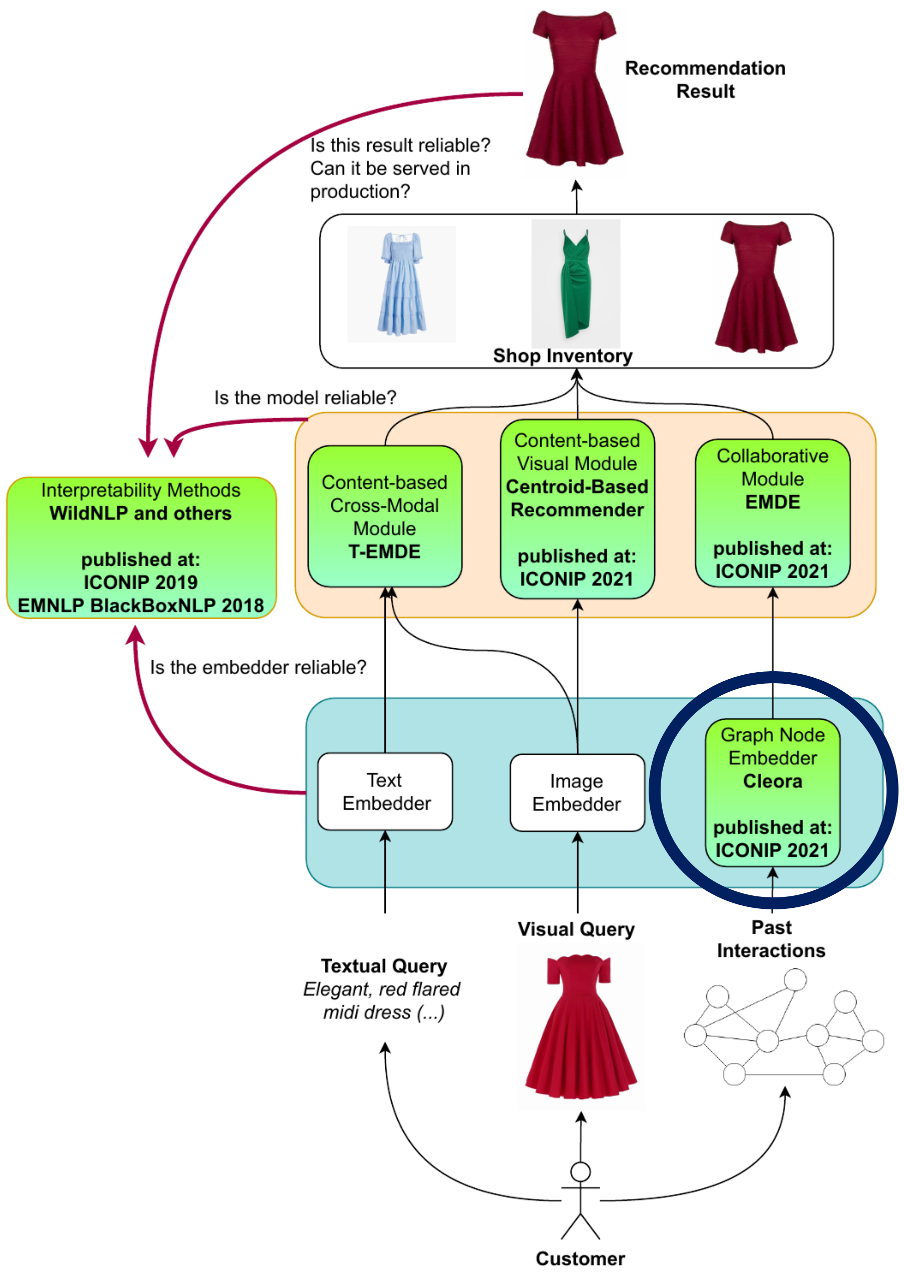


Research Hypotheses

1. Certain forms of graph neural networks (GNNs), usually found in recommender systems, can be simplified by removing weight optimization and non-linearities to obtain a purely unsupervised algorithm based on weighted averaging of node embeddings. Such simplified algorithm can hold high quality of performance. At the same time, it can be much faster to train than a full GNN.
2. Density estimation-based modeling of user item spaces allows to increase the quality in top-k recommendation.
3. Centroid-based image retrieval combined with a neural model allows to increase quality and speed of unimodal image-based content recommendations.
4. Transformation of embeddings from non-matching feature spaces to a unified feature space allows to both increase the quality and speed of cross-modal content recommendation.
5. Increasing the resilience of models against some adversarial examples allows to strengthen the model against related adversarial examples.

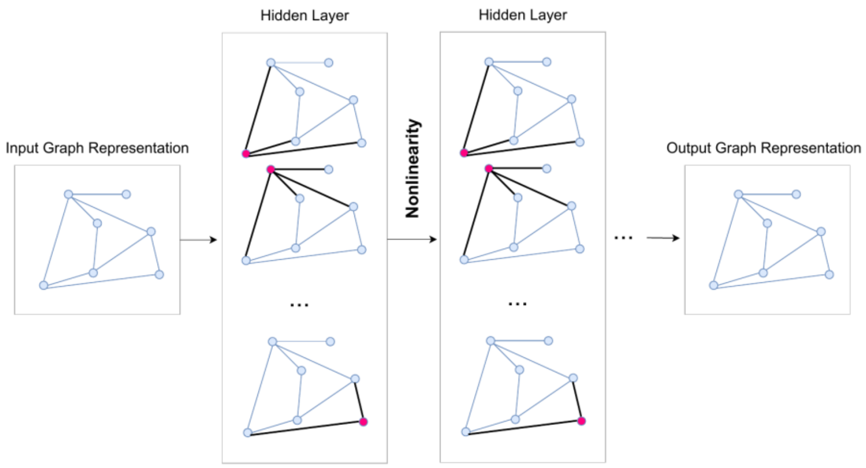
Cleora

- Node embedding algorithm for graph data
- Very fast and scalable
- Inspired by Graph Neural Networks

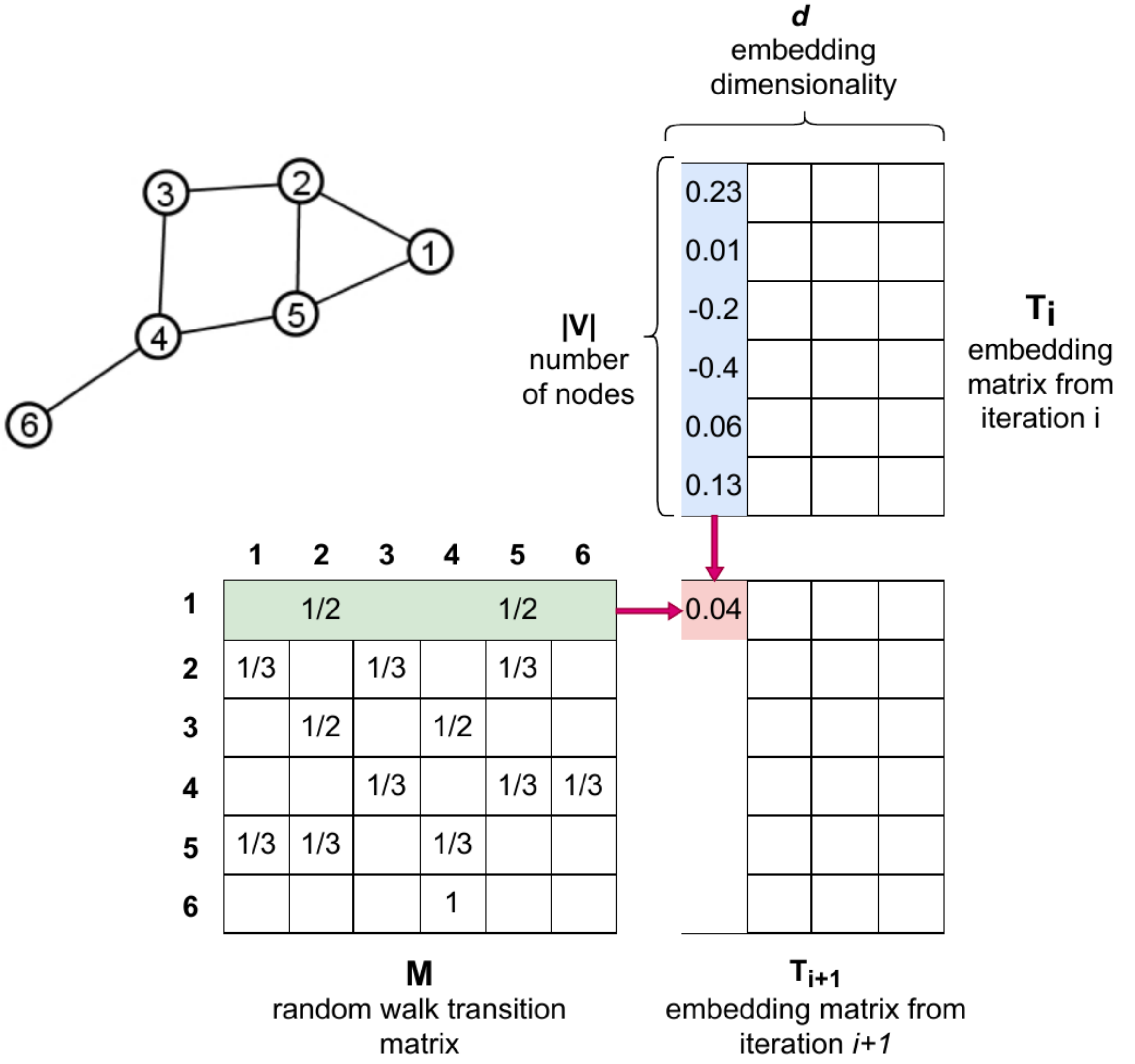


Cleora

Graph Convolutional Network



$$H^{[i+1]} = \sigma(W^{[i]} H^{[i]} A^{*})$$



Cleora: A Simple, Strong and Scalable Graph Embedding Scheme. Barbara Rychalska, Piotr Bąbel. Konrad Gołuchowski, Andrzej Michałowski, Jacek Dąbrowski, Przemysław Biecek. Proceedings of ICONIP 2021

Evidence for this phenomenon also found in:

Simplifying graph convolutional networks. Wu, F., Souza, A., Zhang, T., Fifty, C., Yu, T., & Weinberger, K. (2019, May). International conference on machine learning (pp. 6861-6871). PMLR.

Cleora

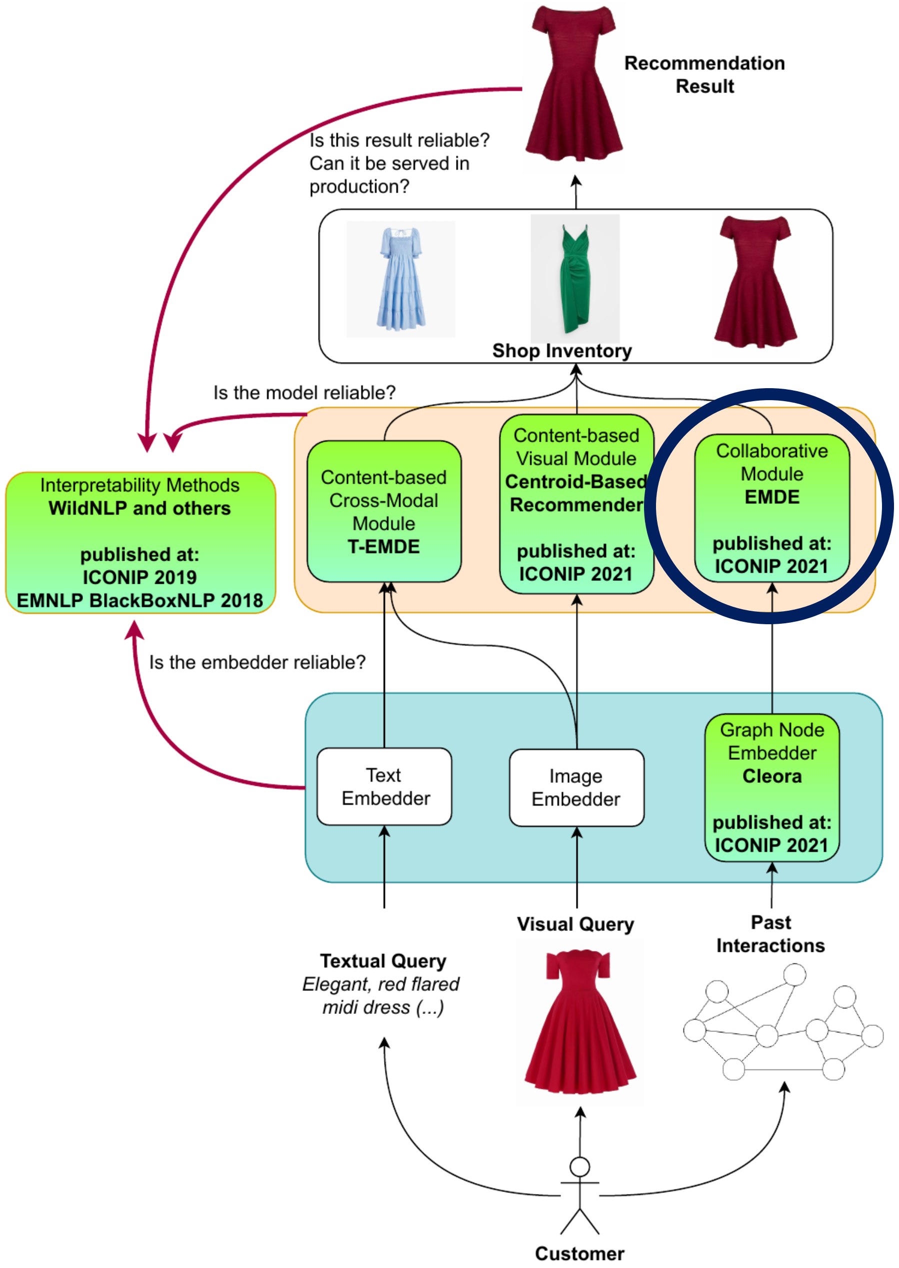
Algorithm	Facebook		YouTube		RoadNet		LiveJournal		Twitter	
	MRR	HR@10	MRR	HR@10	MRR	HR@10	MRR	HR@10	MRR	HR@10
Scalable methods										
Cleora	0.0724	0.1761	0.0471	0.0618	0.9243	0.9429	0.6079	0.6665	0.0355	0.076
PBG	0.0817*	0.2133*	0.0321*	0.0640*	0.8717*	0.9106*	0.5669*	0.6730*	–**	–**
GOSH	0.0924*	0.2319*	0.0280*	0.0590*	0.8756*	0.8977*	0.2242*	0.4012*	–**	–**
Non-scalable methods										
Deepwalk	0.0803*	0.1451*	0.1045*	0.1805*	0.9626*	0.9715*	*timeout*	*timeout*	*timeout*	*timeout*

Algorithm	Cora	CiteSeer	PubMed
AS-GCN	0.874	0.797	0.906
Cleora	0.868	0.757	0.802
AdaGCN	0.855	0.762	0.798
N-GCN	0.830	0.722	0.795
GCN	0.815	0.703	0.790

Cleora: A Simple, Strong and Scalable Graph Embedding Scheme. Barbara Rychalska, Piotr Bąbel. Konrad Gołuchowski, Andrzej Michałowski, Jacek Dąbrowski, Przemysław Biecek. Proceedings of ICONIP 2021

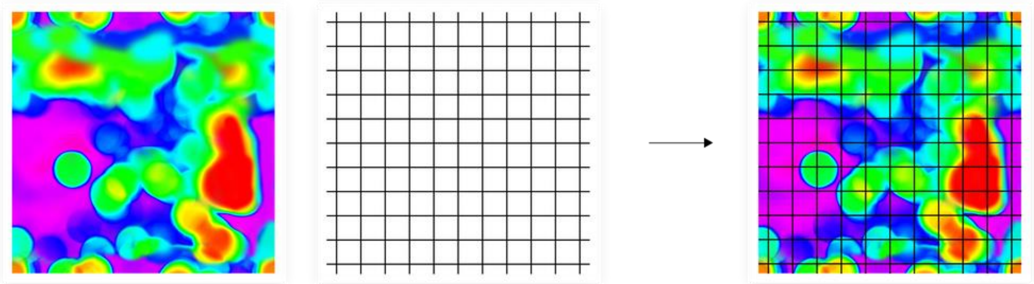
EMDE

- An encoding method which allows to create a user behavioral profile out of items they purchased/viewed/clicked/...
- The encoding very adequately represents sets of items
- Can be used to move the aggregate profile building out of the neural network. E.g. Transformer processes item sequences within the network and it is very time- and resource-consuming. With EMDE we can use a simple feed-forward network.

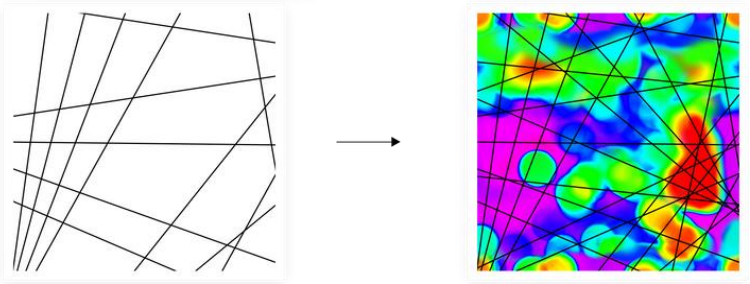


EMDE

Random partitioning of the embedding space



Density-aware partitioning of the embedding space

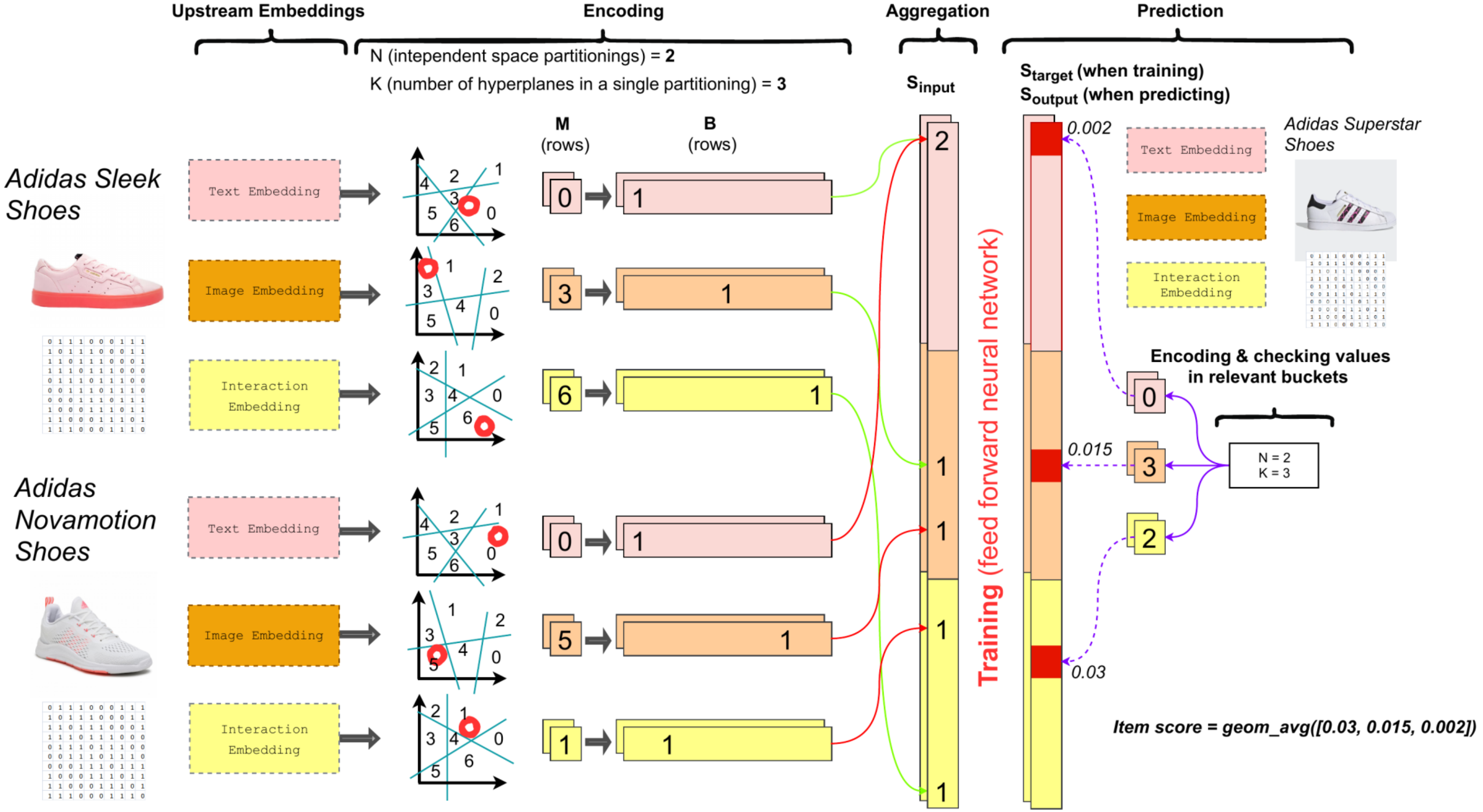


Count-Min Sketch

index	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Hash function 1 →	0	2	0	1	0	0	0	0	0	0	0	0	0	0	0	0
Hash function 2 →	0	0	1	0	1	0	0	1	0	0	0	0	0	0	0	0
Hash function 3 →	0	1	0	0	0	0	0	0	0	1	1	0	0	0	0	0
Hash function 4 →	0	0	1	0	0	0	0	0	0	0	0	2	0	0	0	0

An Efficient Density Estimator for All Recommendation Systems. Jacek Dąbrowski*, Barbara Rychalska*, Michał Daniluk, Dominika Basaj, Konrad Gołuchowski, Piotr Babel, Andrzej Michałowski. * - equal contribution. Proceedings of ICONIP 2021

EMDE



An Efficient Density Estimator for All Recommendation Systems. Jacek Dąbrowski*, Barbara Rychalska*, Michał Daniluk, Dominika Basaj, Konrad Gołuchowski, Piotr Babel, Andrzej Michałowski. * - equal contribution. Proceedings of ICONIP 2021

EMDE

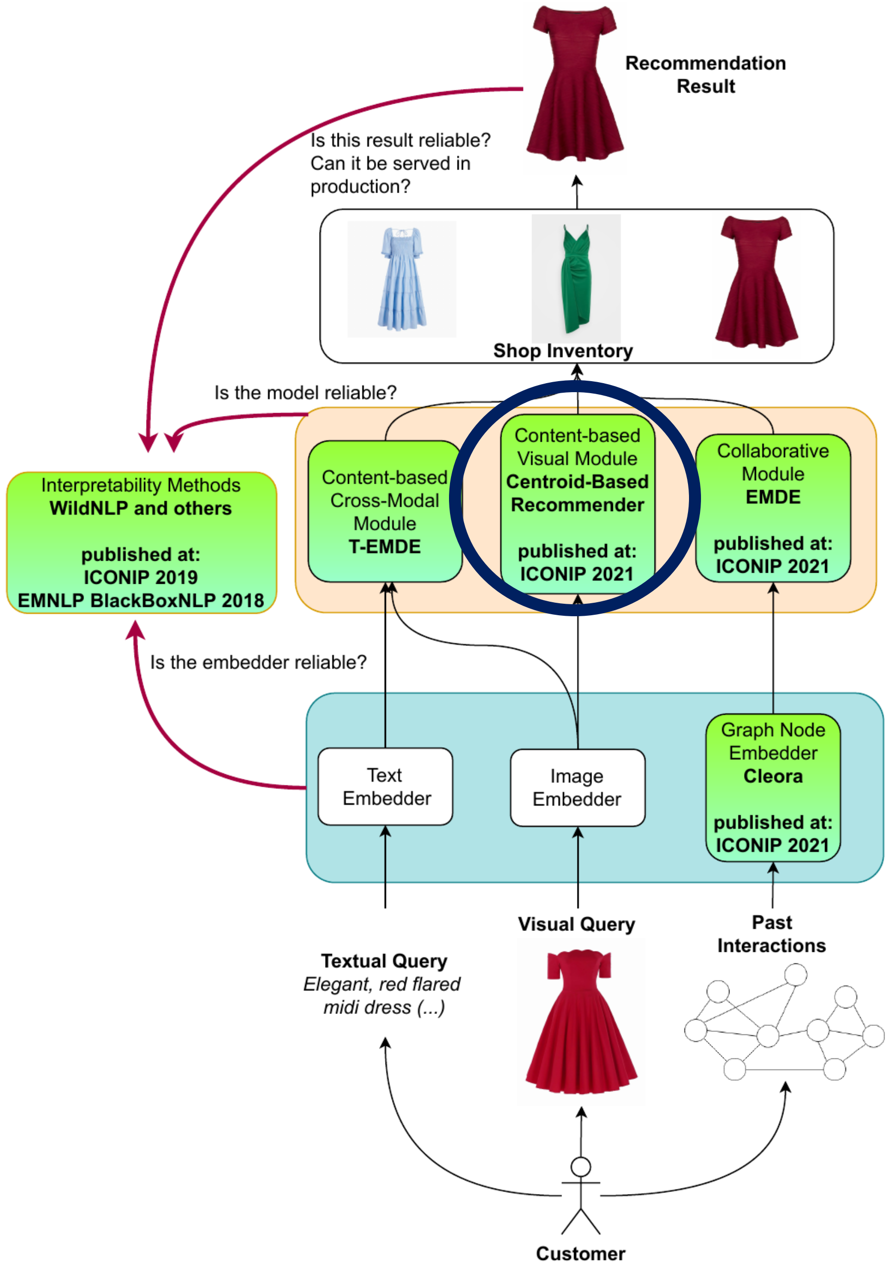
Model	Recall@1	NDCG@5	Recall@5	NDCG@10	Recall@10	NDCG@20	Recall@20
			MovieLens 20M				
EMDE	**0.0529**	**0.2017**	**0.1662**	**0.2535**	**0.2493**	0.3053	0.3523
EASE Steck [2019]	0.0507	0.1990	0.1616	0.2530	0.2465	**0.3078**	**0.3542**
MultVAE Liang et al. [2018]	0.0514	0.1955	0.1627	0.2493	0.2488	0.3052	0.3589
SLIM Ning and Karypis [2011]	0.0474	0.1885	0.1533	0.2389	0.2318	0.2916	0.3350
RP3beta Paudel et al. [2016]	0.0380	0.1550	0.1279	0.2004	0.2007	0.2501	0.3018
			Netflix Prize				
EMDE	**0.0328**	0.1512	0.1101	0.1876	0.1652	0.2332	0.2432
EASE Steck [2019]	0.0323	**0.1558**	**0.1120**	**0.2050**	**0.1782**	**0.2589**	**0.2677**
MultVAE Liang et al. [2018]	0.0313	0.1485	0.1109	0.1957	0.1756	0.2483	0.2645
SLIM Ning and Karypis [2011]	0.0307	0.1484	0.1062	0.1952	0.1688	0.2474	0.2552
RP3beta Paudel et al. [2016]	0.0243	0.0946	0.0595	0.1191	0.0863	0.1578	0.1390

Model	Recall@1	NDCG@5	Recall@5	NDCG@10	Recall@10	NDCG@20	Recall@20
			MovieLens 20M				
EMDE-MM	**0.0670**	**0.2378**	**0.1963**	**0.2890**	**0.2780**	**0.3358**	**0.3710**
EMDE	0.0529	0.2017	0.1662	0.2535	0.2493	0.3053	0.3523
			Netflix Prize				
EMDE-MM	**0.0388**	**0.1574**	**0.1253**	**0.2155**	**0.1875**	**0.2619**	**0.2645**
EMDE	0.0328	0.1512	0.1101	0.1876	0.1652	0.2332	0.2432

An Efficient Density Estimator for All Recommendation Systems. Jacek Dąbrowski*, Barbara Rychalska*, Michał Daniluk, Dominika Basaj, Konrad Gołuchowski, Piotr Babel, Andrzej Michałowski. * - equal contribution. Proceedings of ICONIP 2021

Uni-modal Retrieval

- Uni-modal retrieval is the task of recommendation within a single modality (e.g. pictures) where there is no interaction data, only picture similarity.
- We introduce the concept of item centroid and use it for both training and prediction
- Item centroids:
  - Decrease training time
  - Decrease prediction time
  - Regularize the model and increase metric scores

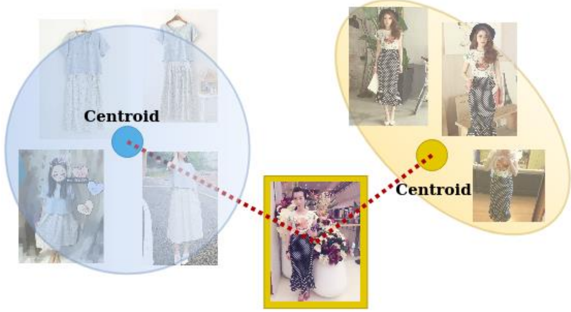


Uni-modal Retrieval

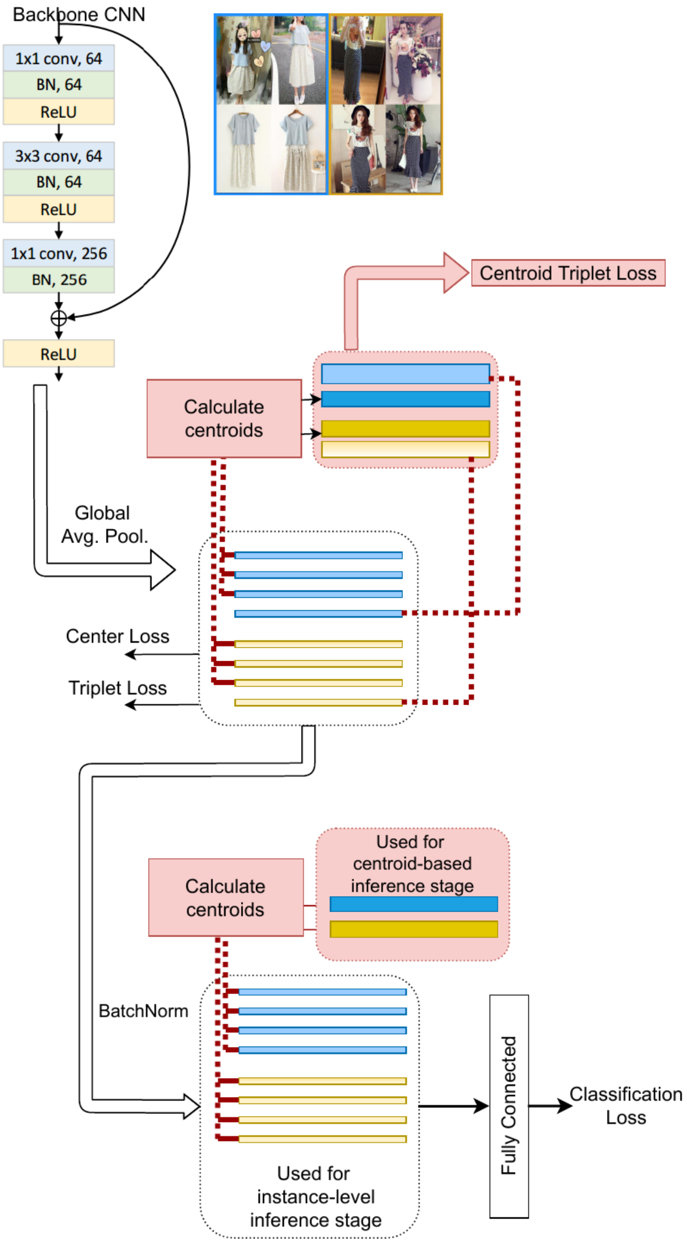
Instance-based Retrieval

$$\mathcal{L}_{triplet} = \left[\|f(A) - f(P)\|_2^2 - \|f(A) - f(N)\|_2^2 + \alpha\right]_+$$

Centroid-based Retrieval



$$\mathcal{L}_{triplet} = \left[\|f(A) - c_P\|_2^2 - \|f(A) - c_N\|_2^2 + \alpha_c\right]_+$$



On the Unreasonable Effectiveness of Centroids in Image Retrieval. Mikołaj Wieczorek*, Barbara Rychalska* and Jacek Dąbrowski. * - equal contribution. Proceedings of ICONIP 2021

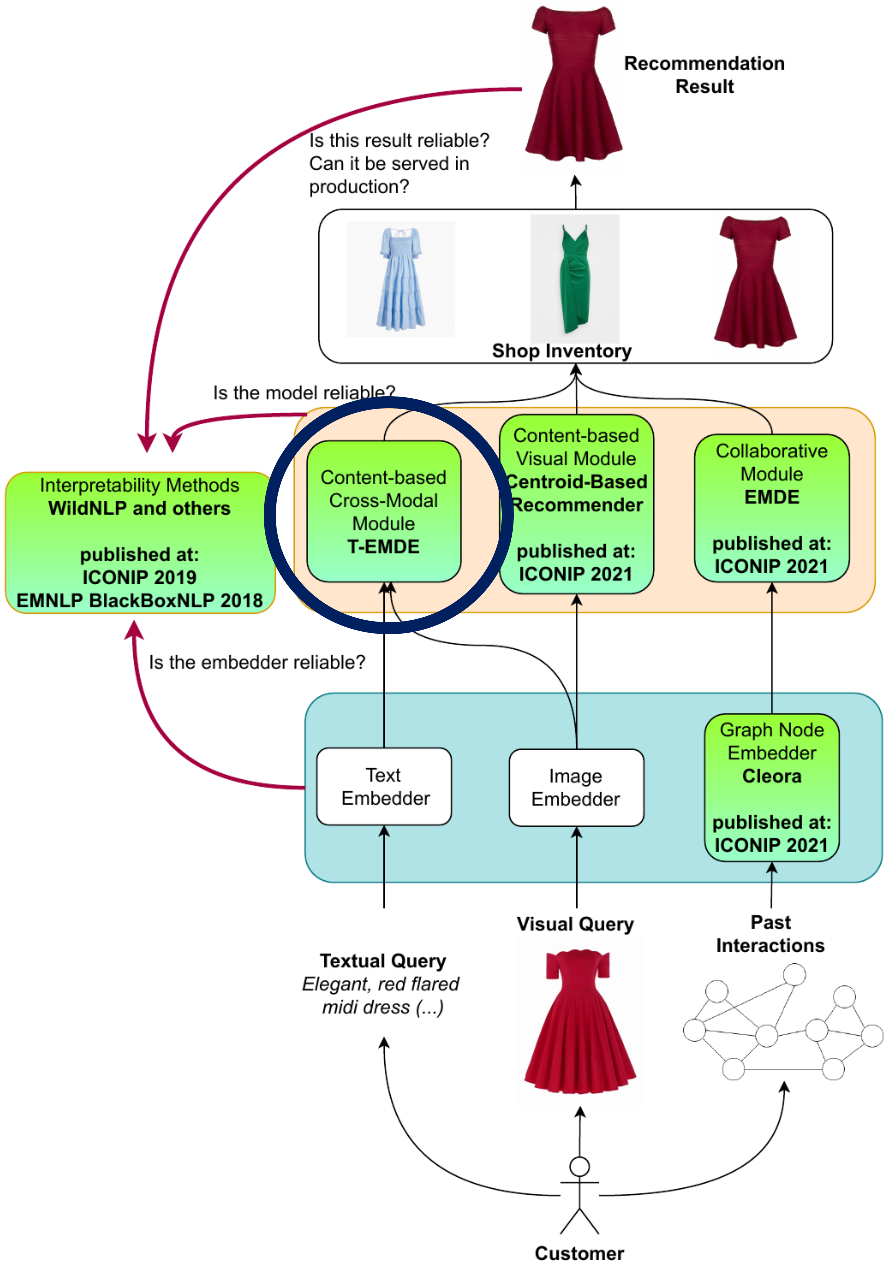
Uni-modal Retrieval

Dataset	Model	mAP	Acc@1	Acc@10	Acc@20	Acc@50
DeepFashion	SOTA (S-R50) [Wieczorek et al., 2020]	0.324	0.281	0.583	0.655	0.742
	CTL-S-R50	0.344	**0.298**	0.612	0.685	0.770
	CTL-S-R50 **CE**	**0.404**	0.294	**0.613**	**0.689**	**0.774**
	SOTA (L-R50IBN) [Wieczorek et al., 2020]	0.430	**0.378**	0.711	0.772	0.841
	CTL-L-R50IBN	0.431	0.376	0.711	0.776	0.847
	CTL-L-R50IBN **CE**	**0.492**	0.373	**0.712**	**0.777**	**0.850**
Street2Shop	SOTA (S-R50) [Wieczorek et al., 2020]	0.320	0.366	0.611	0.606	–
	CTL-S-R50	0.353	0.418	0.594	0.643	0.702
	CTL-S-R50 **CE**	**0.498**	**0.432**	**0.619**	**0.660**	**0.721**
	SOTA (L-R50IBN) [Wieczorek et al., 2020]	0.468	**0.537**	0.698	0.736	–
	CTL-L-R50IBN	0.459	0.533	0.689	0.728	0.782
	CTL-L-R50IBN **CE**	**0.598**	**0.537**	**0.709**	**0.750**	**0.792**

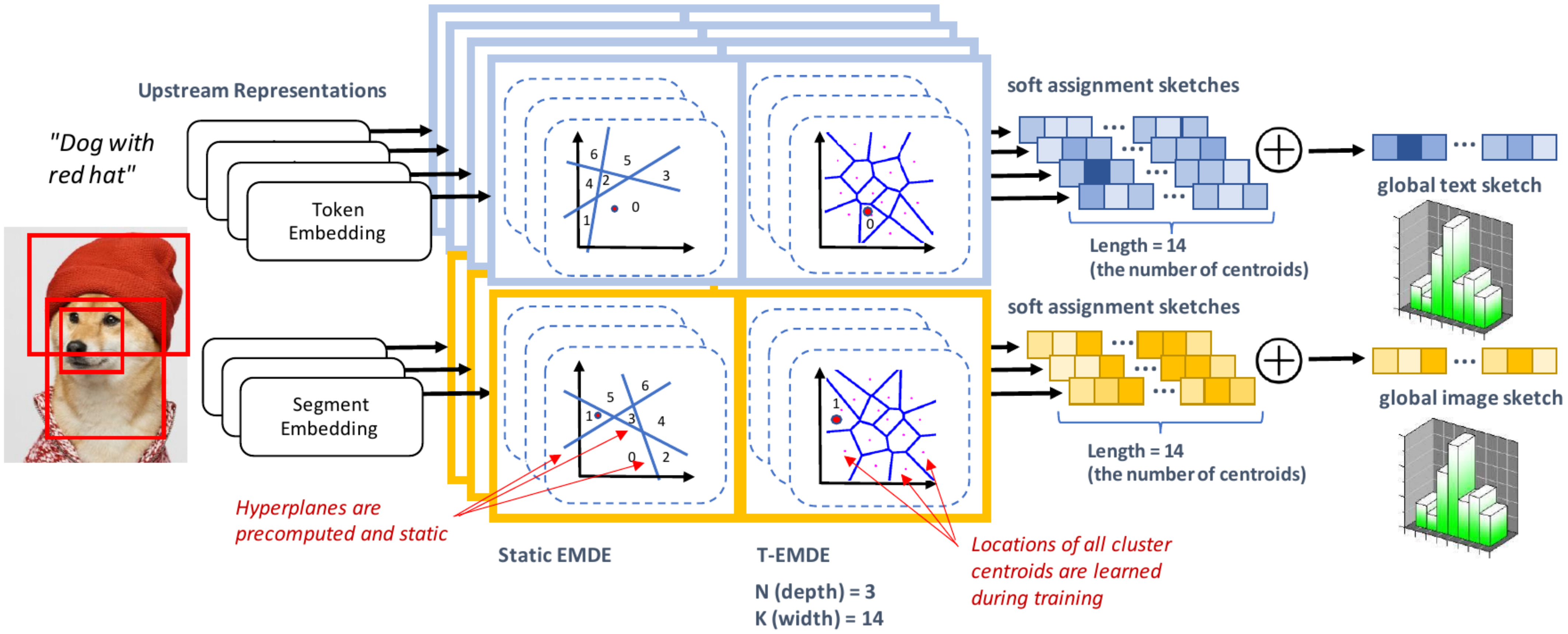
On the Unreasonable Effectiveness of Centroids in Image Retrieval. Mikołaj Wieczorek*, Barbara Rychalska* and Jacek Dąbrowski. * - equal contribution. Proceedings of ICONIP 2021

T-EMDE Multimodal Retrieval

- Multimodal retrieval is the task of recommendation within multiple modalities (e.g. pictures and item descriptions) where there is no interaction data, only picture similarity.
- We introduce the concept of trainable EMDE – movable cluster centroids instead of static item assignments.
- This increases both speed and metric performance.



T-EMDE Multimodal Retrieval



T-EMDE: Sketching-based Global Similarity for Cross-Modal Retrieval. Barbara Rychalska*, Mikołaj Wieczorek*, Jacek Dąbrowski. * - equal contribution. Preprint

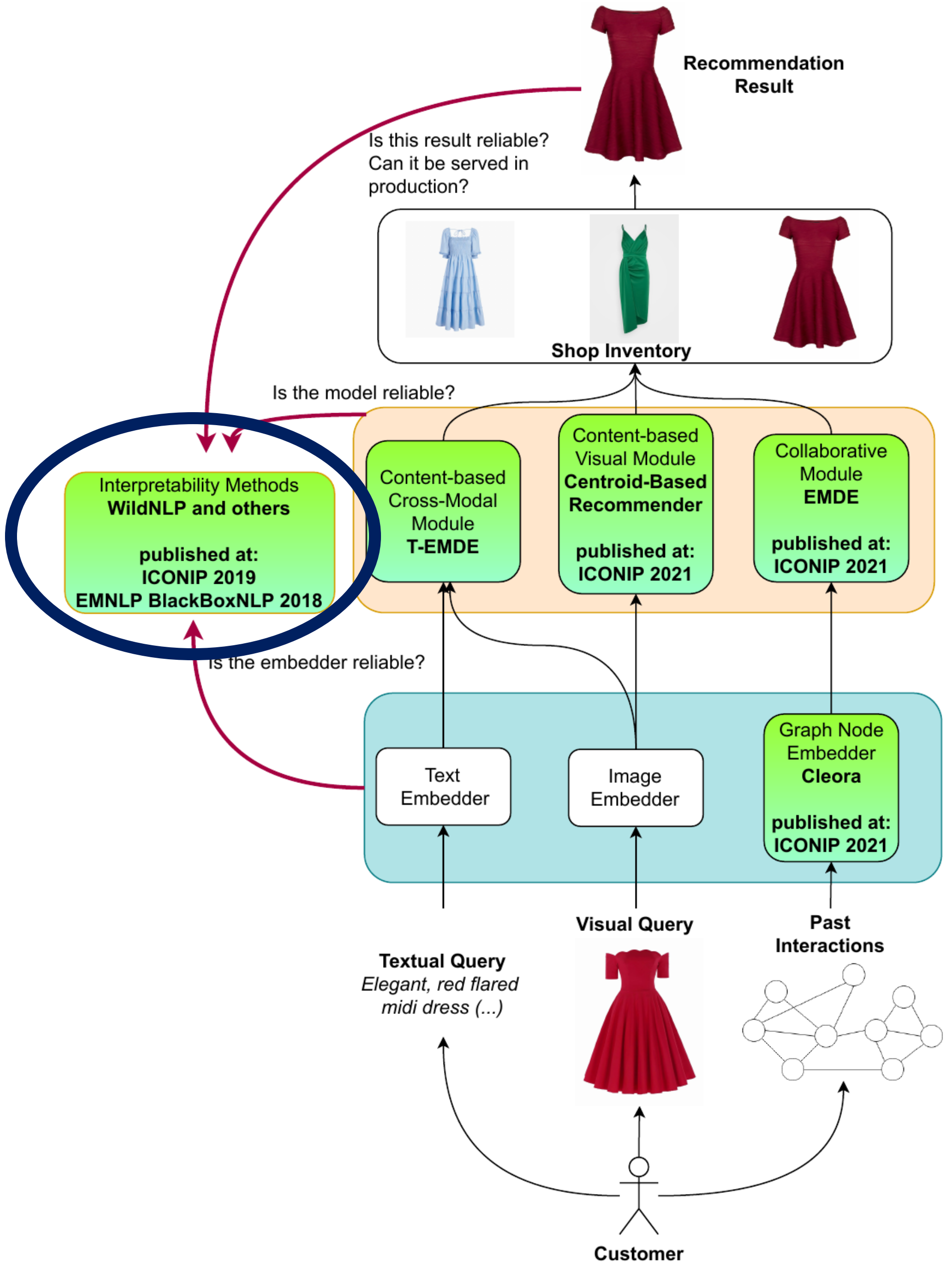
T-EMDE Multimodal Retrieval

Model	Flickr30K dataset							
	Text Retrieval				Image Retrieval			
	R@1	R@5	R@10	MRR	R@1	R@5	R@10	MRR
CAAN Zhang et al. [2020a]	70.1	91.6	97.2	–	52.8	79.0	87.9	–
DP-RNN Chen and Luo [2020]	70.2	91.6	95.8	–	55.5	81.3	88.2	–
PFAN Wang et al. [2019]	70.0	91.8	95.0	–	50.4	78.7	86.1	–
VSRN Li et al. [2019]	71.3	90.6	96.0	–	54.7	81.8	88.2	–
IMRAM Chen et al. [2020a]	74.1	93.0	96.6	–	53.9	79.4	87.2	–
SAF no global	71.1	91.6	96.3	0.803	54.1	80.7	87.3	0.660
SAF reported	73.7	93.3	96.3	–	56.1	81.5	88.0	–
SAF reproduced	74.9	93.9	97.1	0.833	56.3	81.8	88.0	0.676
T-EMDE	75.2	94.2	97.1	0.829	57.1	82.2	88.3	0.682
SGR no global	52.8	83.7	92.8	0.664	49.6	77.3	85.1	0.618
SGR reported	75.2	93.3	96.6	–	56.2	81.0	86.5	–
SGR reproduced	76.3	93.2	96.9	0.835	55.2	80.7	88.0	0.666
T-EMDE	77.5	93.1	97.2	0.845	56.9	82.0	87.5	0.679
SGRAF reported	78.4	**94.6**	**97.5**	–	58.2	83.0	89.1	–
SGRAF reproduced	77.5	94.4	97.0	0.849	58.4	83.2	89.0	0.693
T-EMDE	**78.8**	94.4	**97.5**	**0.858**	**59.6**	**83.6**	**89.2**	**0.702**

T-EMDE: Sketching-based Global Similarity for Cross-Modal Retrieval. Barbara Rychalska*, Mikołaj Wieczorek*, Jacek Dąbrowski. * - equal contribution. Preprint

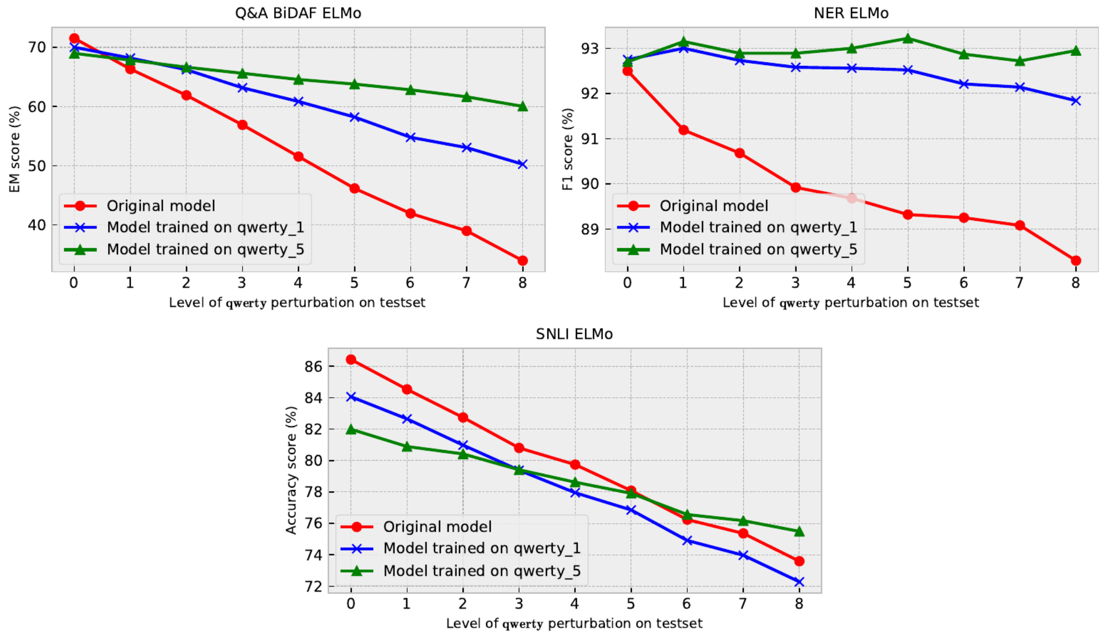
Interpretability & Robustness

- Recommender systems in production need to be monitored for vulnerabilities (such as other production systems).
- We introduce WildNLP - a method for detecting vulnerabilities and fixing them by adversarial training.



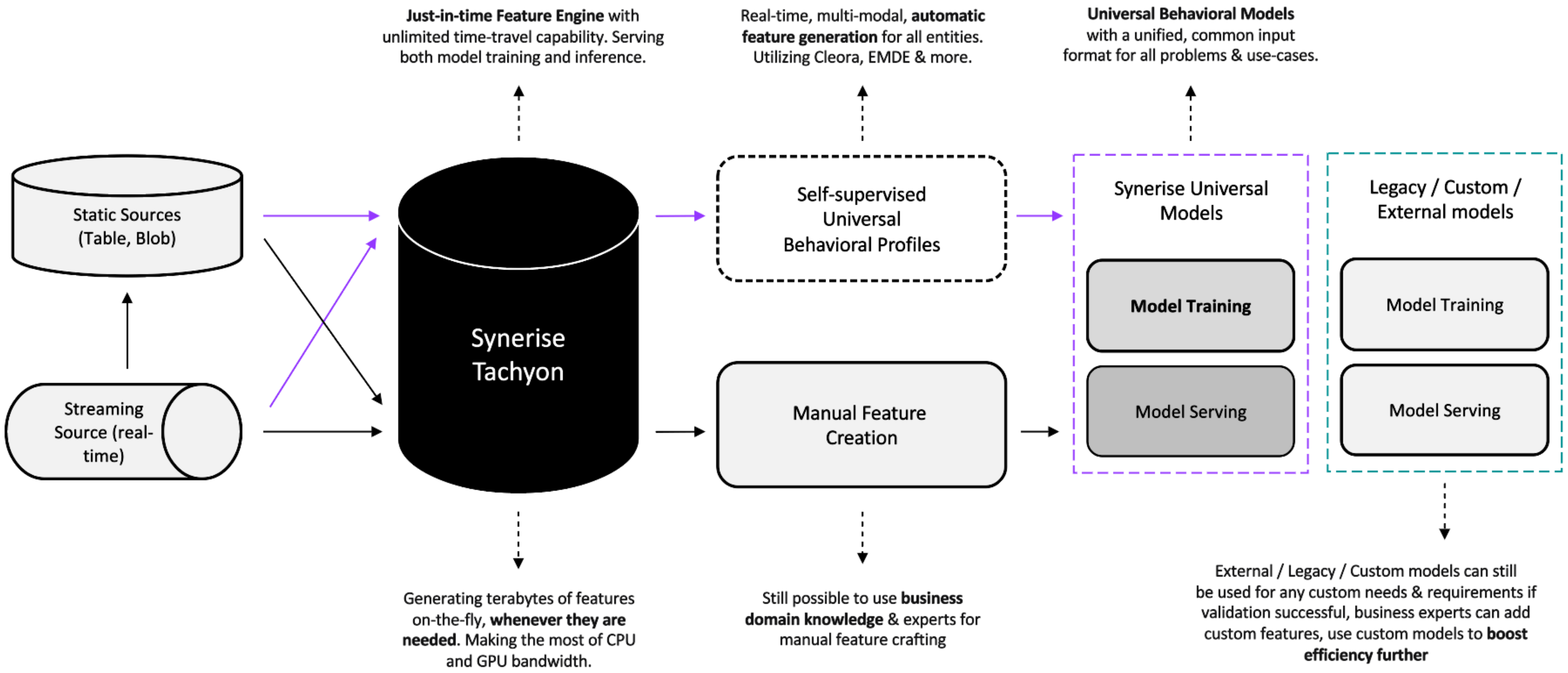
Interpretability & Robustness

Aspect	Example sentence
Original	Warsaw was believed to be one of the most beautiful cities in the world.
Article	Warsaw was believed to be one of **a** most beautiful cities in world.
Swap	Warsaw **aws** believed to be one **fo teh** most beautiful cities in the world.
Qwerty	Wa**d**saw was b**d**lieved to be one of the most beautiful citi**e**e in the world.
Remove_char	Warsaw was believed to be one **o th** most **eautiful** cities in the world.
Remove_space	Warsaw was believed **tobe** one of the most beautiful cities in the world.
Original	You cannot accidentally commit vandalism. It used to be a rare occurrence.
Misspelling	You can not **accidentaly** commit vandalism. It used to be a rare **occurrance**.
Original	Bus Stops for Route 6, 6.1
Digits2words	Bus Stops for Route **six**, **six point one**
Original	Choosing between affect and effect can be scary.
Homophones	Choosing between **effect** and effect can **bee** scary.
Original	Laughably foolish or false: an absurd explanation.
Negatives	**Laughab*y fo*lish** or **fal*e**: an **a*surd** explanation.
Original	Sometimes it is good to be first, and sometimes it is good to be last.
Positives	Sometimes it is **go*d** to be first, and sometimes it is **goo*** to be last.
Marks	Sometimes, it is good to be first and sometimes, it, is good to be last.



Models in the Wild: On Corruption Robustness of Neural NLP Systems. Barbara Rychalska, Dominika Basaj, Alicja Gosiewska, Przemysław Biecek. Proceedings of ICONIP 2019

Commercialization



Synerise Monad - Real-Time Multimodal Behavioral Modeling. Jacek Dąbrowski, Barbara Rychalska. Proceedings of the 31st ACM International Conference on Information and Knowledge Management (CIKM'22)

Synerise Monad – A Foundation Model for Behavioral Event Data. Barbara Rychalska, Szymon Łukasik, Jacek Dąbrowski. Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'23)

Bibliography

- *Cleora: A Simple, Strong and Scalable Graph Embedding Scheme.* Barbara Rychalska, Piotr Bąbel. Konrad Gołuchowski, Andrzej Michałowski, Jacek Dąbrowski, Przemysław Biecek. Proceedings of ICONIP 2021 (140 MNiSW points).
- *An Efficient Density Estimator for All Recommendation Systems.* Jacek Dąbrowski*, Barbara Rychalska*, Michał Daniluk, Dominika Basaj, Konrad Gołuchowski, Piotr Babel, Andrzej Michałowski. * - equal contribution. Proceedings of ICONIP 2021 (140 MNiSW points).
- *On the Unreasonable Effectiveness of Centroids in Image Retrieval.* Mikołaj Wieczorek*, Barbara Rychalska* and Jacek Dąbrowski. * - equal contribution. Proceedings of ICONIP 2021 (140 MNiSW points).
- *T-EMDE: Sketching-based Global Similarity for Cross-Modal Retrieval.* Barbara Rychalska*, Mikołaj Wieczorek*, Jacek Dąbrowski. * - equal contribution. Preprint.
- *Models in the Wild: On Corruption Robustness of Neural NLP Systems.* Barbara Rychalska, Dominika Basaj, Alicja Gosiewska, Przemysław Biecek. Proceedings of ICONIP 2019 (140 MNiSW points).
- *Synerise Monad - Real-Time Multimodal Behavioral Modeling.* Jacek Dąbrowski, Barbara Rychalska. Proceedings of the 31st ACM International Conference on Information and Knowledge Management (CIKM ’22) (140 MNiSW points)
- *Synerise Monad – A Foundation Model for Behavioral Event Data.* Barbara Rychalska, Szymon Łukasik, Jacek Dąbrowski. Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR ’23) (200 MNiSW points)