Yandi LI (Boris)

& (+86) 18401596880 | ⊠ liyandi007@163.com | ♀ yandili | 🖬 Yandi | ႙ yandi.space

◊ Beijing, China

Work Experience

Sina Weibo 🔊	Beijing, China	11/2014 - Present					
• Machine Learning Algorithm Engin	eer						
 Developed user profiling and interest prediction models Improved short-term interest prediction from daily to real-time (within 10 minutes) 							
Yahoo! Research	Barcelona, Spain	02/2014 - 07/2014					
• Research Intern							
 Published paper on applying space 	syntax to online mapping tools						
Education							
UPC & Université Lumière Lyon 2	Barcelona, Spain & Lyon, France	09/2012 - 09/2014					
• Master's in Data Mining and Know	vledge Management (DMKM)						
• Erasmus Mundus Scholarship							
 Key courses: Kernel-based Learning 	g, Statistical NLP, Advanced Statistical Modeling						
Zhejiang University	Hangzhou, China	09/2008 - 06/2012					
• Bachelor's in Mathematics and Ap	plied Mathematics						
• GPA: 3.87/4.0, Top 5%							
UCLA	Los Angeles, USA	08/2009 - 09/2009					
• Summer School							
 Courses in Finance and English Wr 	 Courses in Finance and English Writing 						
Publications							
[P1] Apply Space Syntax to Online Ma	apping Tools, WSDM 2017						
• Yandi Li , Nicola Barbieri (Tumblr), D	aniele Quercia (Bell Labs)						
• Key Technologies: Factorization Mach	ine, BPR, PostgreSQL, PostGIS						
[P2] Different Implementations and Co	omparisons of the Chebyshev-Tao Method						
• Undergraduate thesis, [Part of the The	esis], [Literature Review], [Defense]						
Projects							
Weibo User Modeling		05/2022 - Present					
• Developed Deep Learning Models for I	User Interest						
• Denoising VAE, Caser structure, a	udience normalization, unified feature embedding						
• Build Comprehensive user profiles							
• Key Technologies: Fuxi-CTR, PyTore	ch, Ollama, Huggingface						
Weibo User Interest	Technical Lead	05/2019 - 05/2022					

• Long and short-term interest retrieval

- Improved real-time interest updates, integrated more behavior logs
- Challenges: sensitivity, Matthew effect, granularity issues (Patent CN115827966A)

о	Ad targeting: transitio	ned from	statistical	models to	supervised	algorithms	(GBDT),	improving	targeting
	conversion rates								
0	• Key Technologies: Flink, Hivemall, LightGBM, Grafana, Streamlit, Clickhouse								
lm	age Recommendation	5		Technical L	ead			08/2017 -	04/2019

08/2016 - 10/2017

03/2015 - 11/2016

- Implemented end-to-end image recommendation
- Developed models for image semantic vectors using contrastive learning
- Key Technologies: asyncio, sanic, PyTorch, Spark, faiss

Image Feature Extraction

- Developed image classification and face recognition
- Fraud detection: identify fake celebrity avatars
- Intelligent image cropping for thumbnails
- Key Technologies: TensorFlow, Keras, Docker, CNN

Recommendation Content Repository

- Built text classification(TextCNN using title, summary, body, author dimension, multi-model stacking) and clickbait detection models
- Similar image deduplication system
- Key Technologies: Keras, scikit-learn, Elasticsearch, MySQL, phash

Personal Projects

KDD Cup (Authorship Disambiguati	on) Lyon, France	03/2013 - 06/2013			
 Disambiguated author names in a large academic database, [Report] 					
 Multilingual text processing, LDA topic extraction 					
\circ Key Technologies: R, PostgreSQL, Python, $ at MT_EX $					
Yet Another Datalog Interpreter	Lyon, France	09/2012 - 06/2013			
 Key Technologies: Datalog, Ocaml, SQL, [Report] 					
Optimal Rescue Search Path	Hangzhou, China	2010 - 2012			
• Key Technologies: Graph theory, Hamilton path, [Report]					
Skills and Interests					

• Programming Languages: Python, SQL, R, Matlab, Java

• Tools and Frameworks: PyTorch, Keras, scikit-learn, Flask, ELK, Flink, Docker

• Hobbies: Tennis
 , Ping Pong
 , Tai Chi€, Skiing
 , DIY